Assessment Title
Scenario-Based Activity Design and Implementation (AWS)

CCC602| Assessment-4

The Student Name | Geni Bahar
Student ID: 27085022

Tutor/Accessor | Ovesh Vohra

Programme
Diploma in Cloud Computing and Cyber Security (120 Credits)

Course
CCC602: Data Engineering in the Public Cloud
(Level 6, 30 Credits)

Date: 12th February 2026

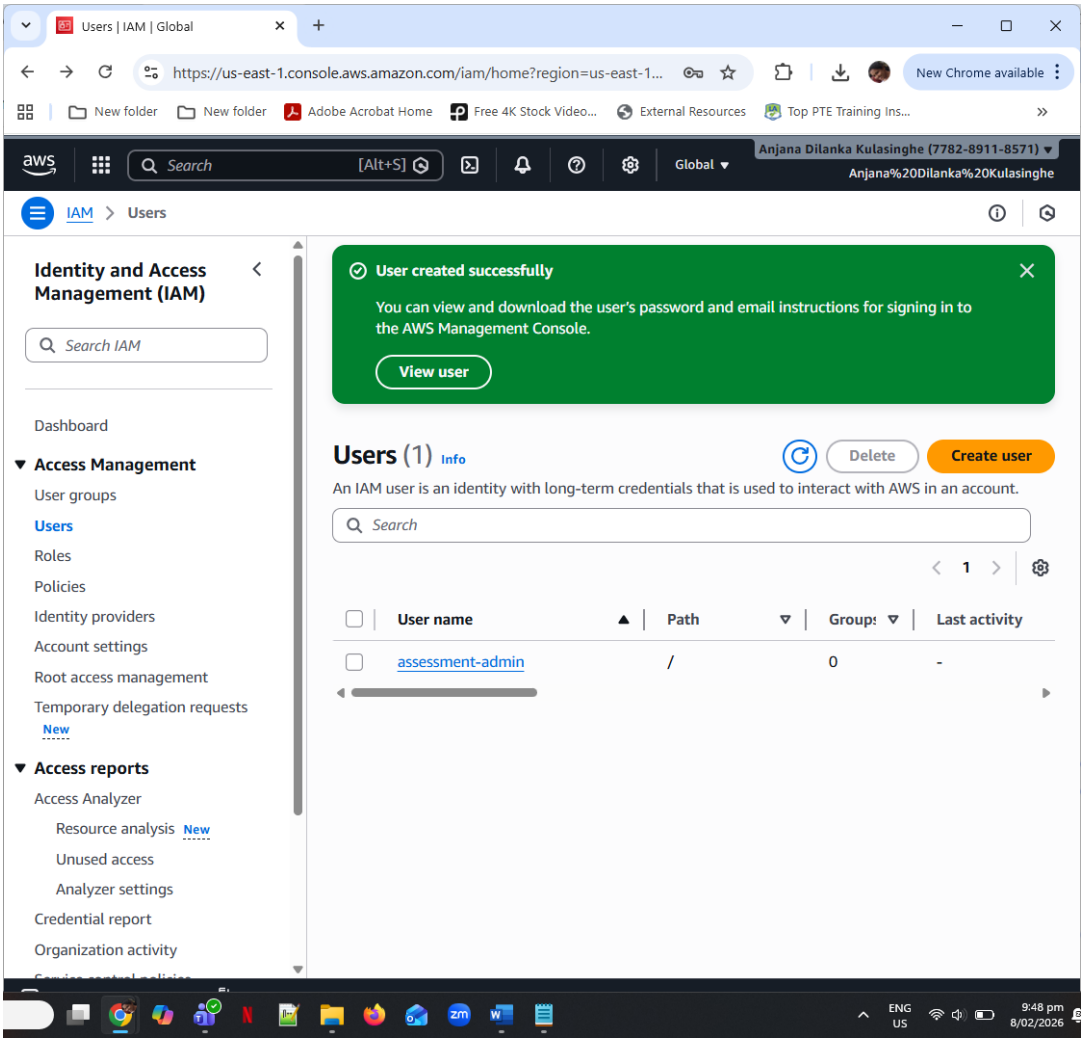
Table of Contents

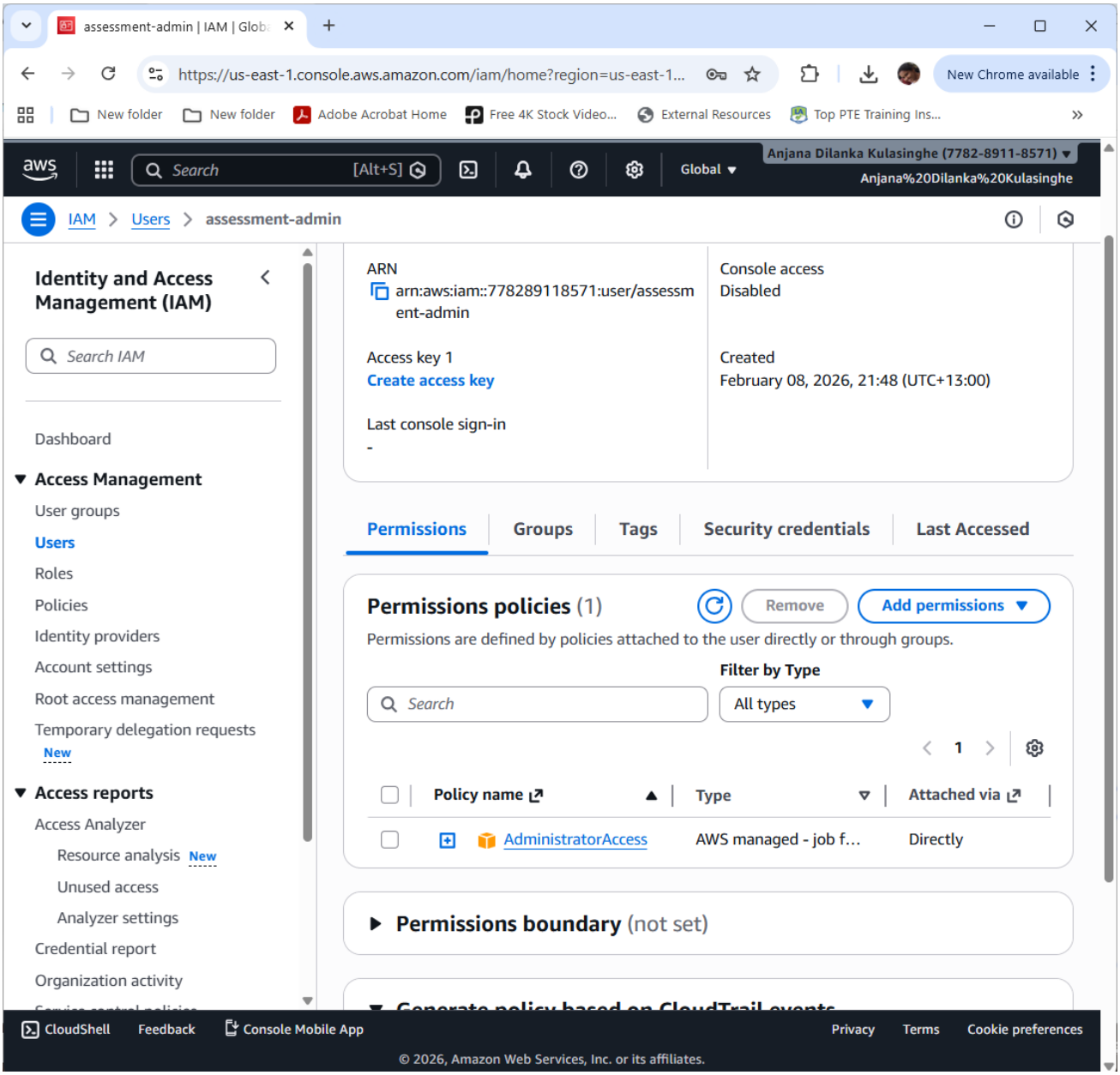
Task 1: Cloud Infrastructure Setup and Automation for Data Ingestion, Cataloguing, and Real-Time Notifications Using AWS

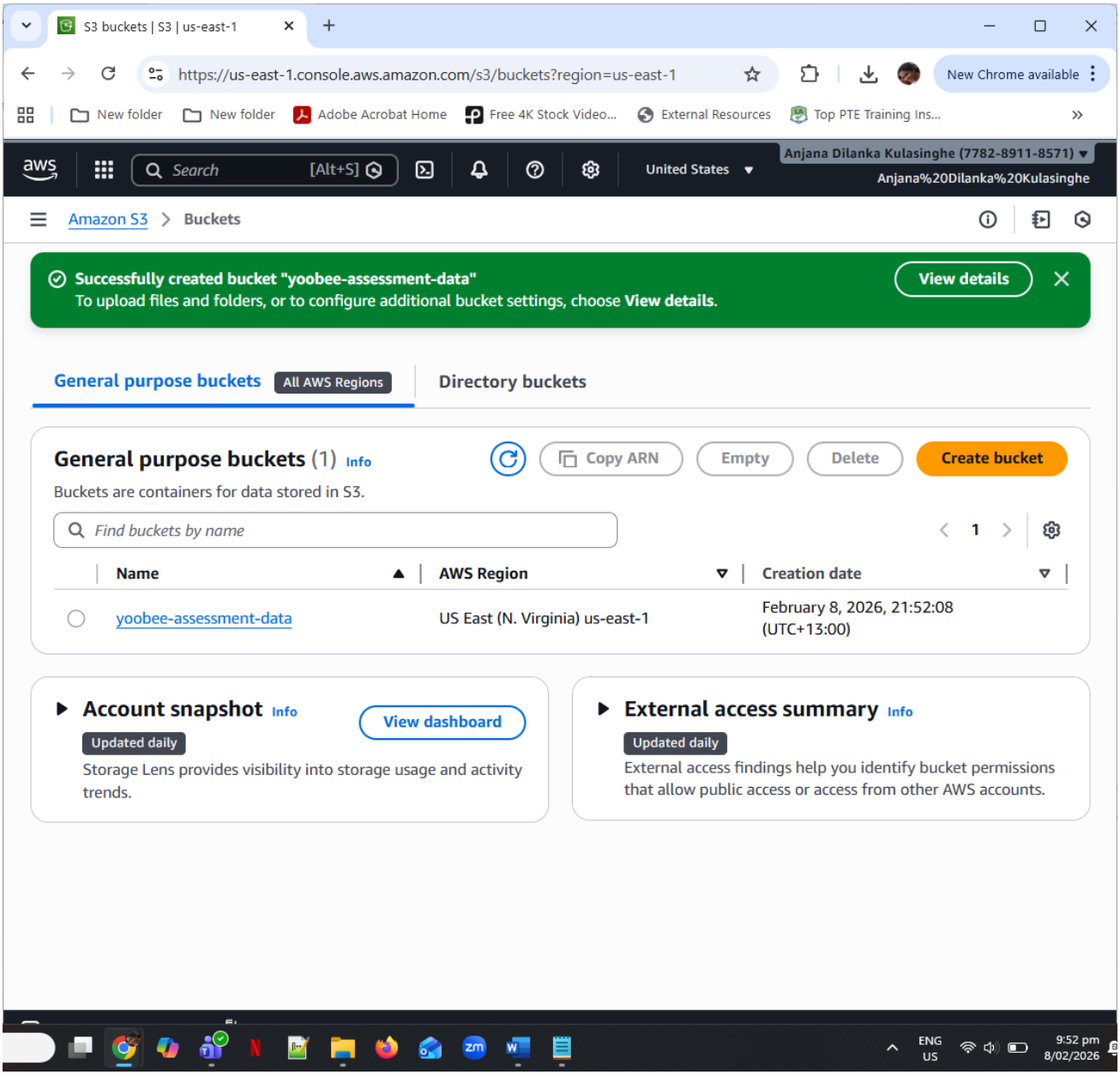
In this task, I configured the core AWS services required to build the foundation of the automated data pipeline. The main objective was to securely upload datasets into Amazon S3, catalogue them using AWS Glue, and implement real-time monitoring using Lambda and SNS. This step established the storage, security, and automation components required for later querying and visualisation.

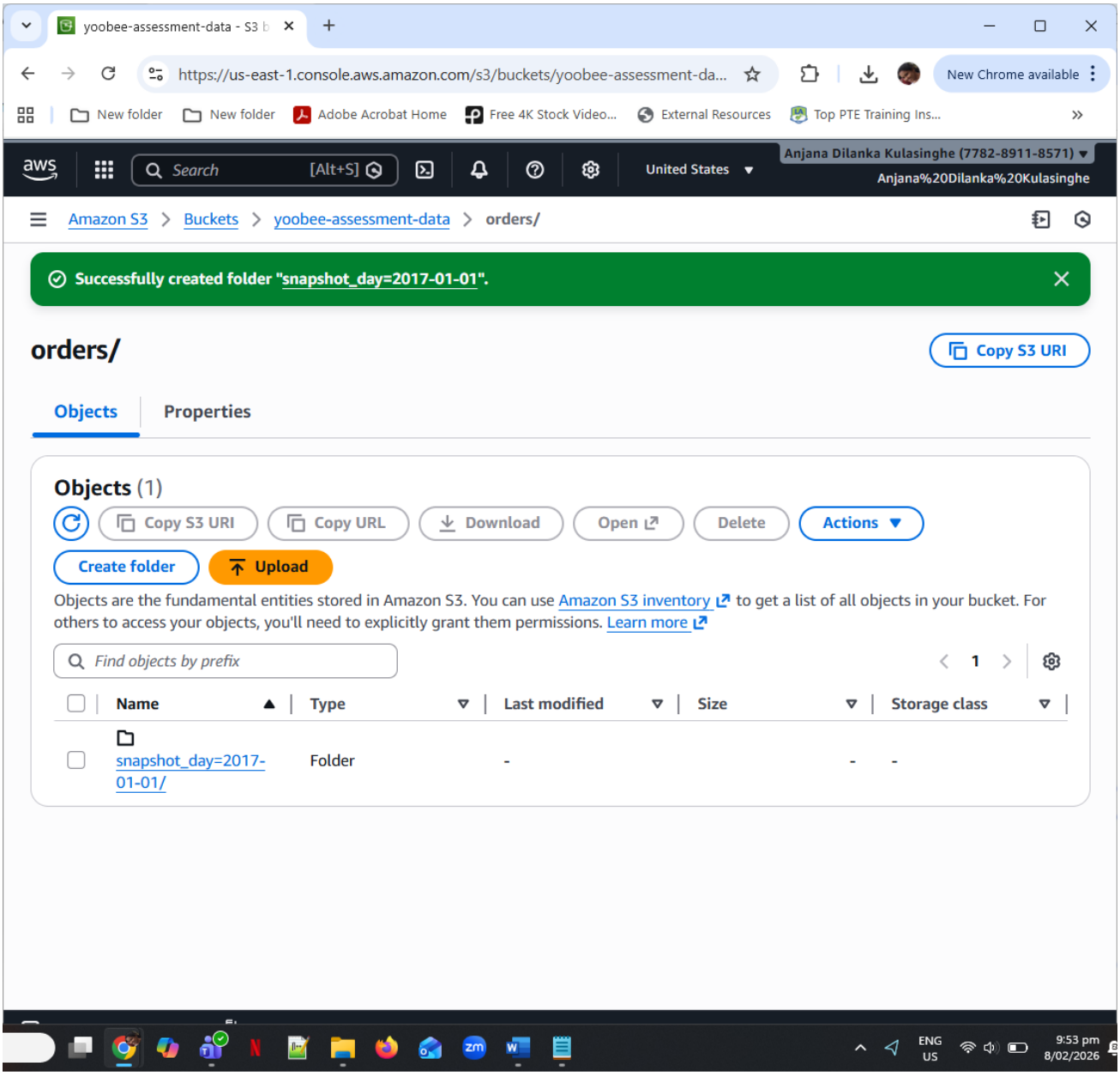
Activities: 1.1: IAM and S3 Setup

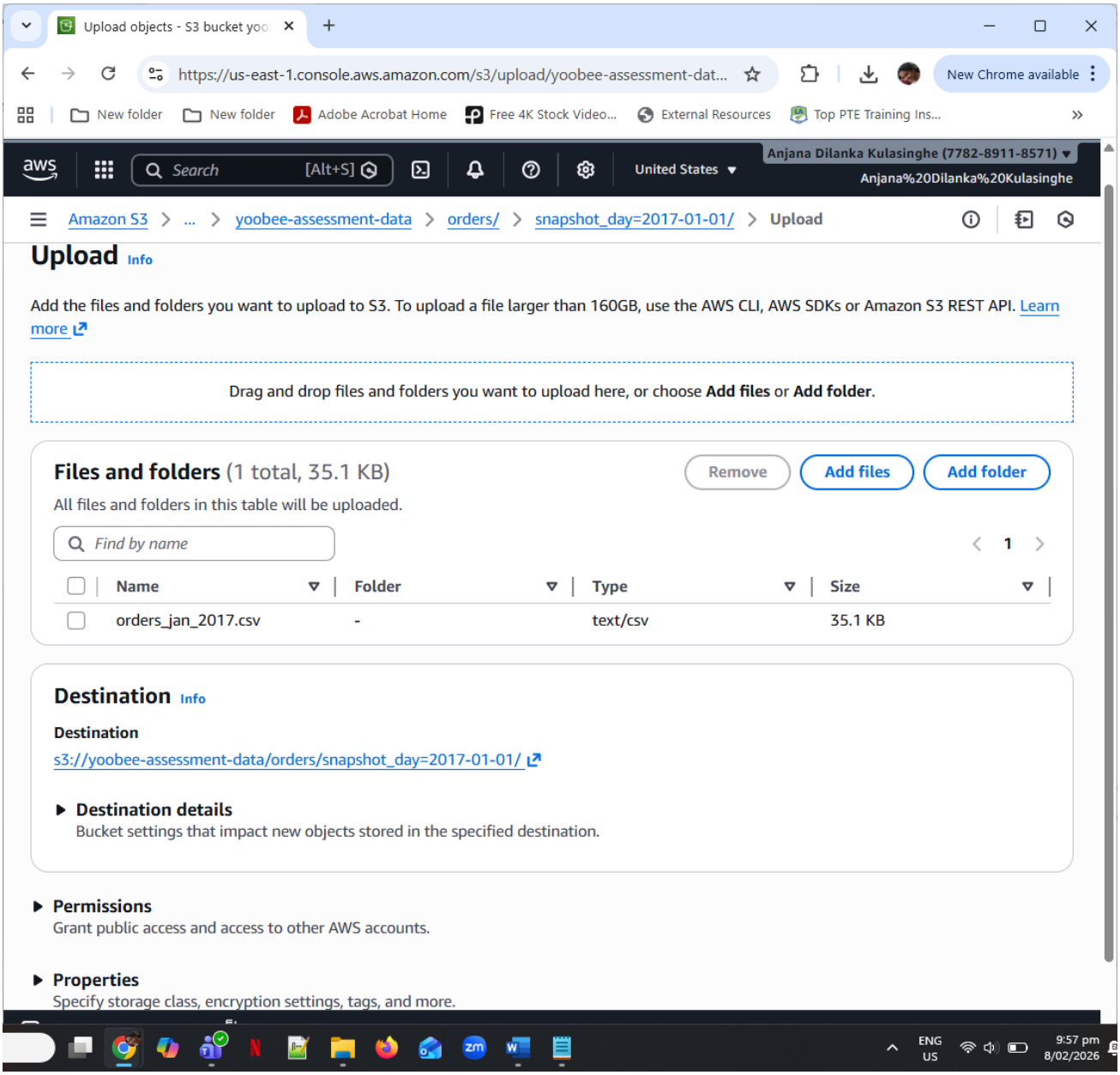
An IAM user was created with AdministratorAccess permissions for the purpose of this assessment environment. Although in a real enterprise scenario, least-privilege access would be applied, full access is simplified configuration during the learning process.
An S3 bucket was created and structured using partition-style folders such as /orders/snapshot_day=2017-01-01. This structure is important because AWS Glue detects partitions based on folder naming conventions. The January 2017 CSV dataset was uploaded into the correct snapshot folder to simulate raw data ingestion.

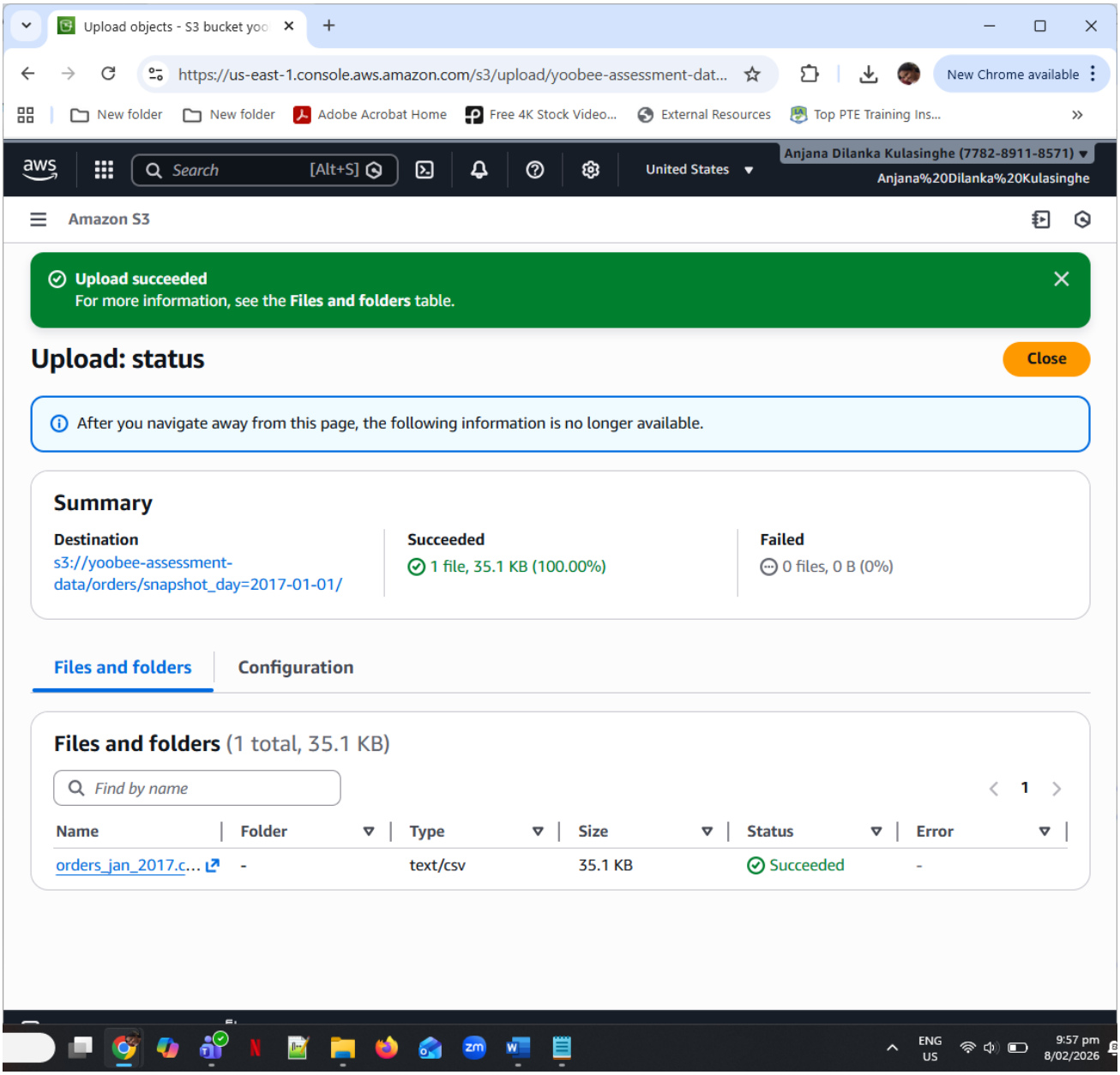



assessment-admin | IAM | Glob
https://us-east-1.console.aws.amazon.com/iam/home?region=us-east-1...
New Chrome available
New folder
New folder
Adobe Acrobat Home
Free 4K Stock Video...
External Resources
Top PTE Training Ins...
aws
Search
[Alt+S]
Global
Anjana Dilanka Kulasinghe (7782-8911-8571)
Anjana%20Dilanka%20Kulasinghe
IAM > Users > assessment-admin
Identity and Access Management (IAM)
Search IAM
Dashboard
Access Management
User groups
Users
Roles
Policies
Identity providers
Account settings
Root access management
Temporary delegation requests
New
Access reports
Access Analyzer
Resource analysis New
Unused access
Analyzer settings
Credential report
Organization activity
ARN
arn:aws:iam::778289118571:user/assessment-admin
Access key 1
Create access key
Last console sign-in
-
Console access
Disabled
Created
February 08, 2026, 21:48 (UTC+13:00)
Permissions
Groups
Tags
Security credentials
Last Accessed
Permissions policies (1)
Remove
Add permissions
Permissions are defined by policies attached to the user directly or through groups.
Filter by Type
Search
All types
1
Policy name
Type
Attached via
AdministratorAccess
AWS managed - job f...
Directly
Permissions boundary (not set)
CloudShell
Feedback
Console Mobile App
Privacy
Terms
Cookie preferences
© 2026, Amazon Web Services, Inc. or its affiliates.


S3 buckets | S3 | us-east-1
https://us-east-1.console.aws.amazon.com/s3/buckets?region=us-east-1
New Chrome available
New folder
New folder
Adobe Acrobat Home
Free 4K Stock Video...
External Resources
Top PTE Training Ins...
aws
Search
[Alt+S]
United States
Anjana Dilanka Kulasinghe (7782-8911-8571)
Anjana%20Dilanka%20Kulasinghe
Amazon S3
Buckets
Successfully created bucket "yoobee-assessment-data"
To upload files and folders, or to configure additional bucket settings, choose View details.
View details
General purpose buckets
All AWS Regions
Directory buckets
General purpose buckets (1) Info
Buckets are containers for data stored in S3.
Copy ARN
Empty
Delete
Create bucket
Find buckets by name
1
Name
AWS Region
Creation date
yoobee-assessment-data
US East (N. Virginia) us-east-1
February 8, 2026, 21:52:08 (UTC+13:00)
Account snapshot Info
Updated daily
Storage Lens provides visibility into storage usage and activity trends.
View dashboard
External access summary Info
Updated daily
External access findings help you identify bucket permissions that allow public access or access from other AWS accounts.
ENG US
9:52 pm
8/02/2026

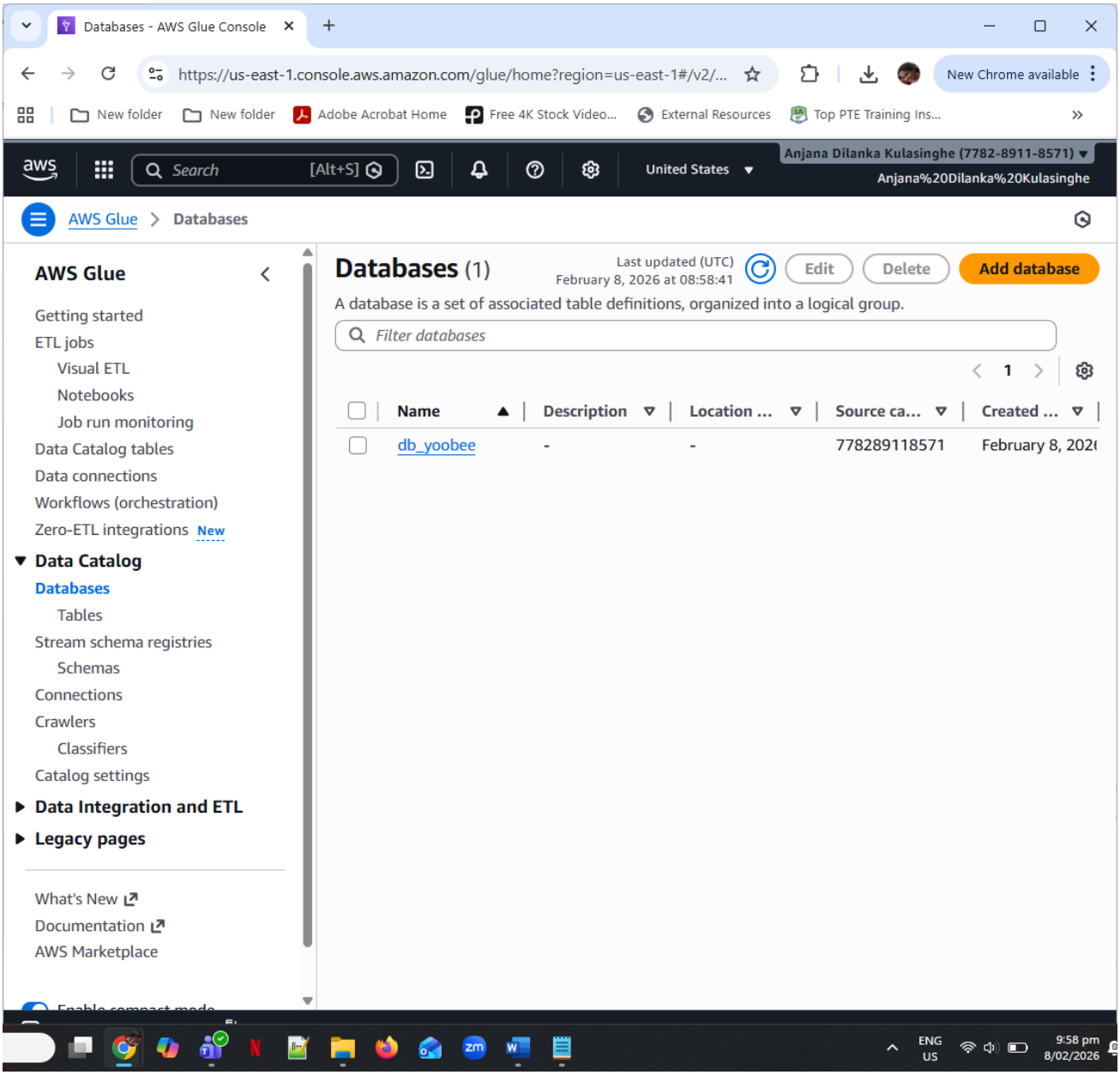

yoobee-assessment-data - S3
https://us-east-1.console.aws.amazon.com/s3/buckets/yoobee-assessment-da...
New Chrome available
New folder
New folder
Adobe Acrobat Home
Free 4K Stock Video...
External Resources
Top PTE Training Ins...
aws
Search
[Alt+S]
United States
Anjana Dilanka Kulasinghe (7782-8911-8571)
Anjana%20Dilanka%20Kulasinghe
Amazon S3 > Buckets > yoobee-assessment-data > orders/
Successfully created folder "snapshot_day=2017-01-01".
orders/
Copy S3 URI
Objects
Properties
Objects (1)
Copy S3 URI
Copy URL
Download
Open
Delete
Actions
Create folder
Upload
Objects are the fundamental entities stored in Amazon S3. You can use Amazon S3 inventory to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. Learn more
Find objects by prefix
1
Name
Type
Last modified
Size
Storage class
snapshot_day=2017-01-01/
Folder
-
-
-
ENG
US
9:53 pm
8/02/2026

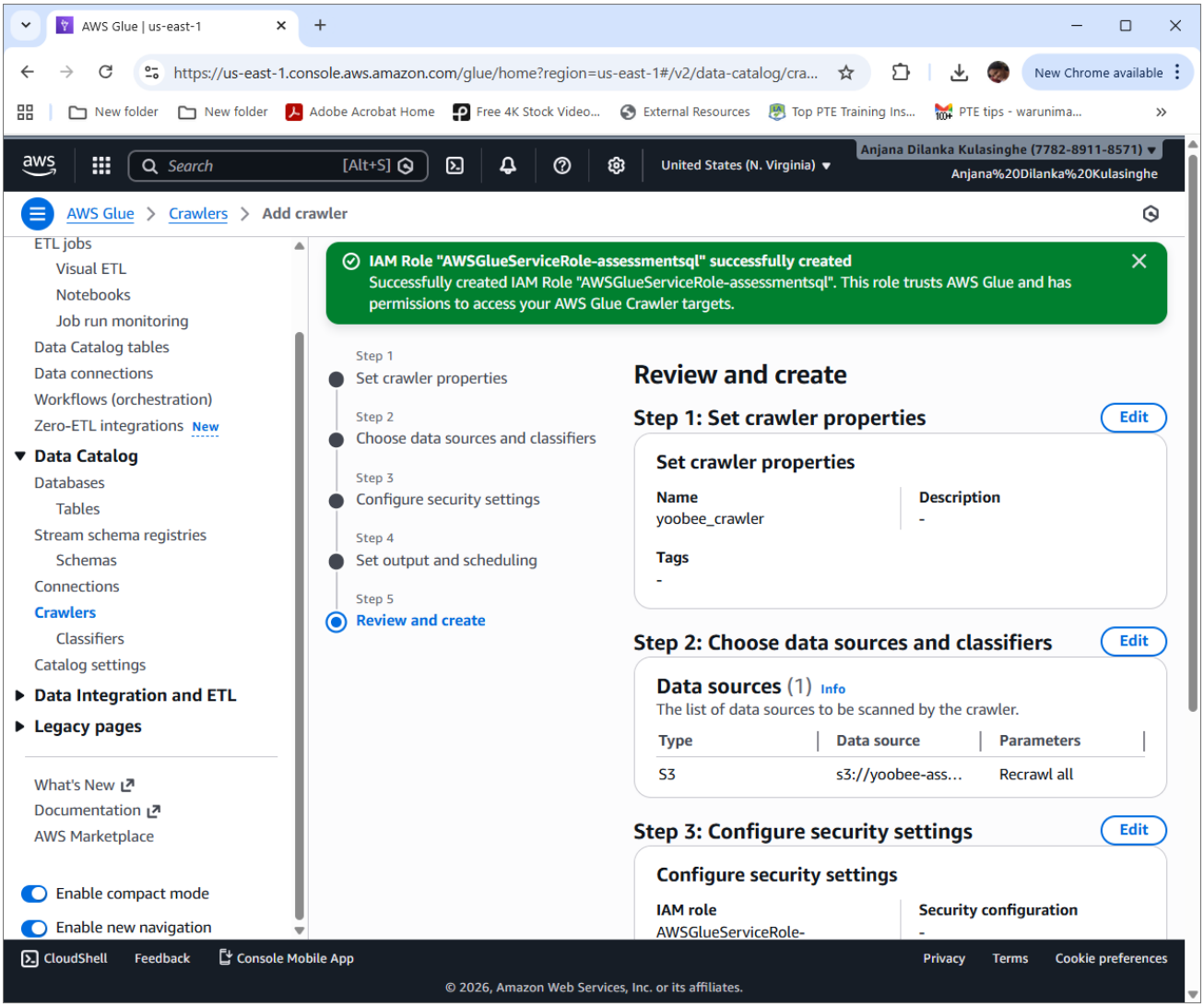

Upload objects - S3 bucket yoo
https://us-east-1.console.aws.amazon.com/s3/upload/yoobee-assessment-dat...
New Chrome available
New folder
New folder
Adobe Acrobat Home
Free 4K Stock Video...
External Resources
Top PTE Training Ins...
aws
Search
[Alt+S]
United States
Anjana Dilanka Kulasinghe (7782-8911-8571)
Anjana%20Dilanka%20Kulasinghe
Amazon S3
...
yoobee-assessment-data
orders/
snapshot_day=2017-01-01/
Upload
Upload Info
Add the files and folders you want to upload to S3. To upload a file larger than 160GB, use the AWS CLI, AWS SDKs or Amazon S3 REST API. Learn more
Drag and drop files and folders you want to upload here, or choose Add files or Add folder.
Files and folders (1 total, 35.1 KB)
Remove
Add files
Add folder
All files and folders in this table will be uploaded.
Find by name
1
Name
Folder
Type
Size
orders_jan_2017.csv
-
text/csv
35.1 KB
Destination Info
Destination
s3://yoobee-assessment-data/orders/snapshot_day=2017-01-01/
Destination details
Bucket settings that impact new objects stored in the specified destination.
Permissions
Grant public access and access to other AWS accounts.
Properties
Specify storage class, encryption settings, tags, and more.
ENG
US
9:57 pm
8/02/2026

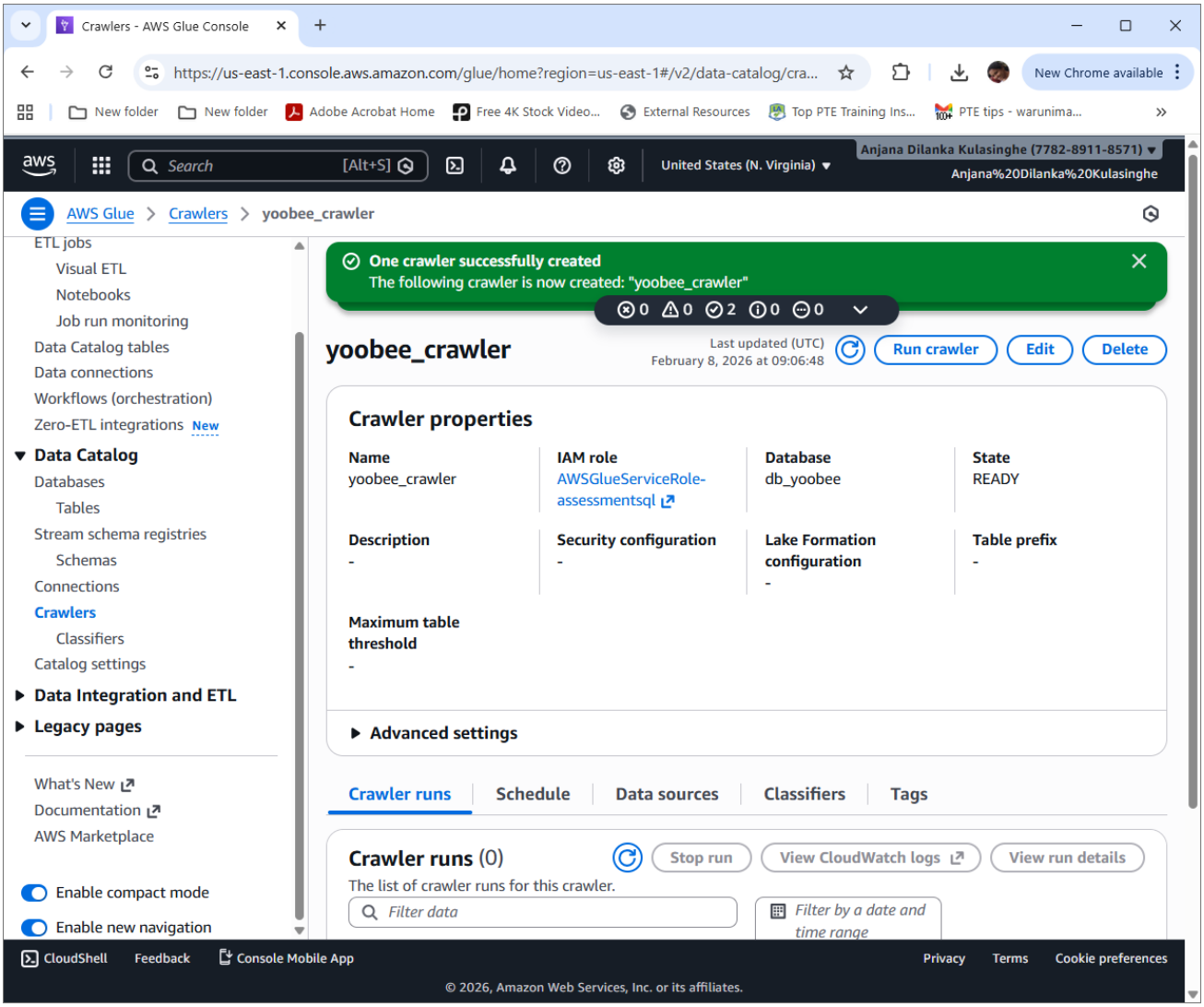


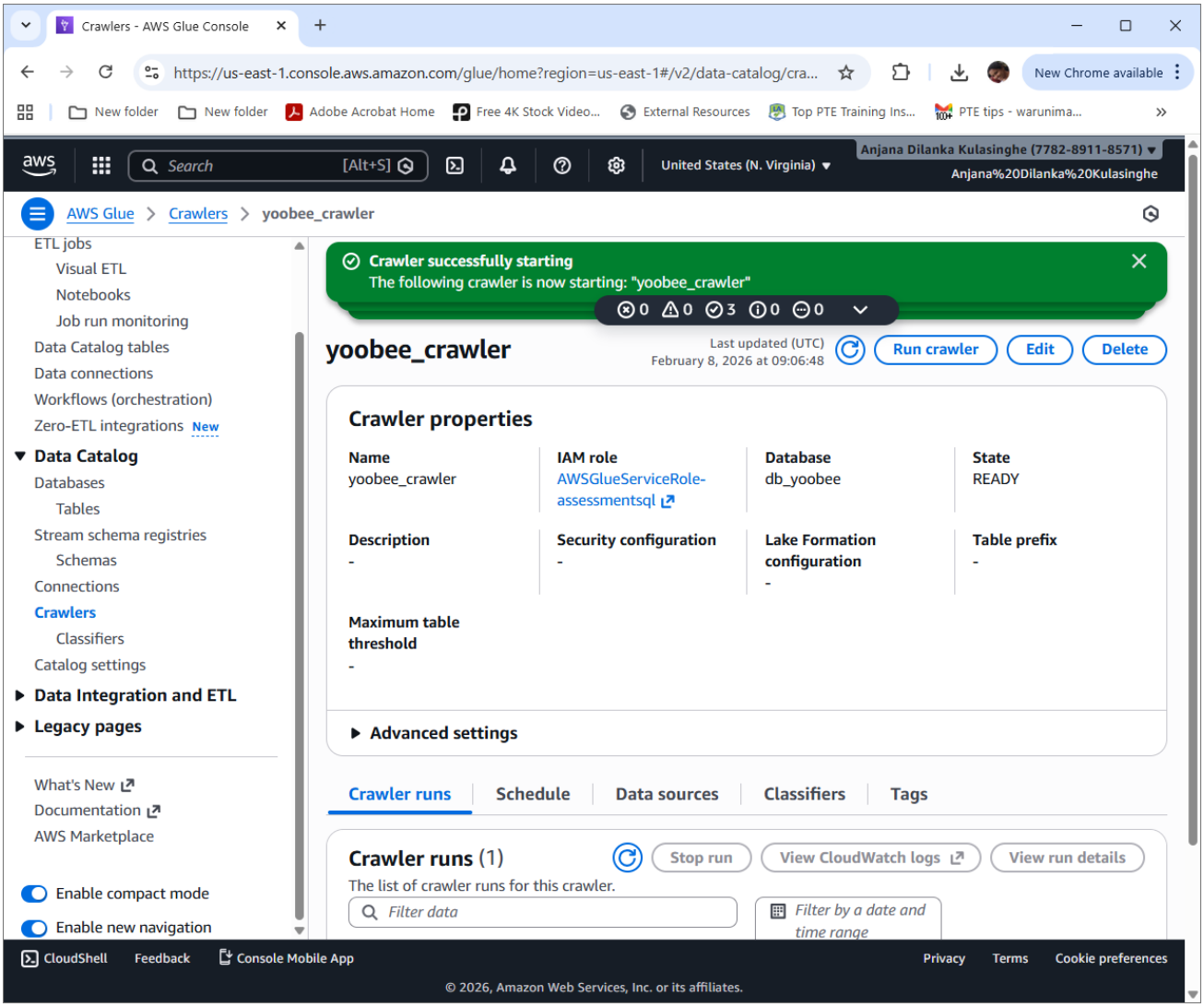
Activities: 1.2: AWS Glue Configuration

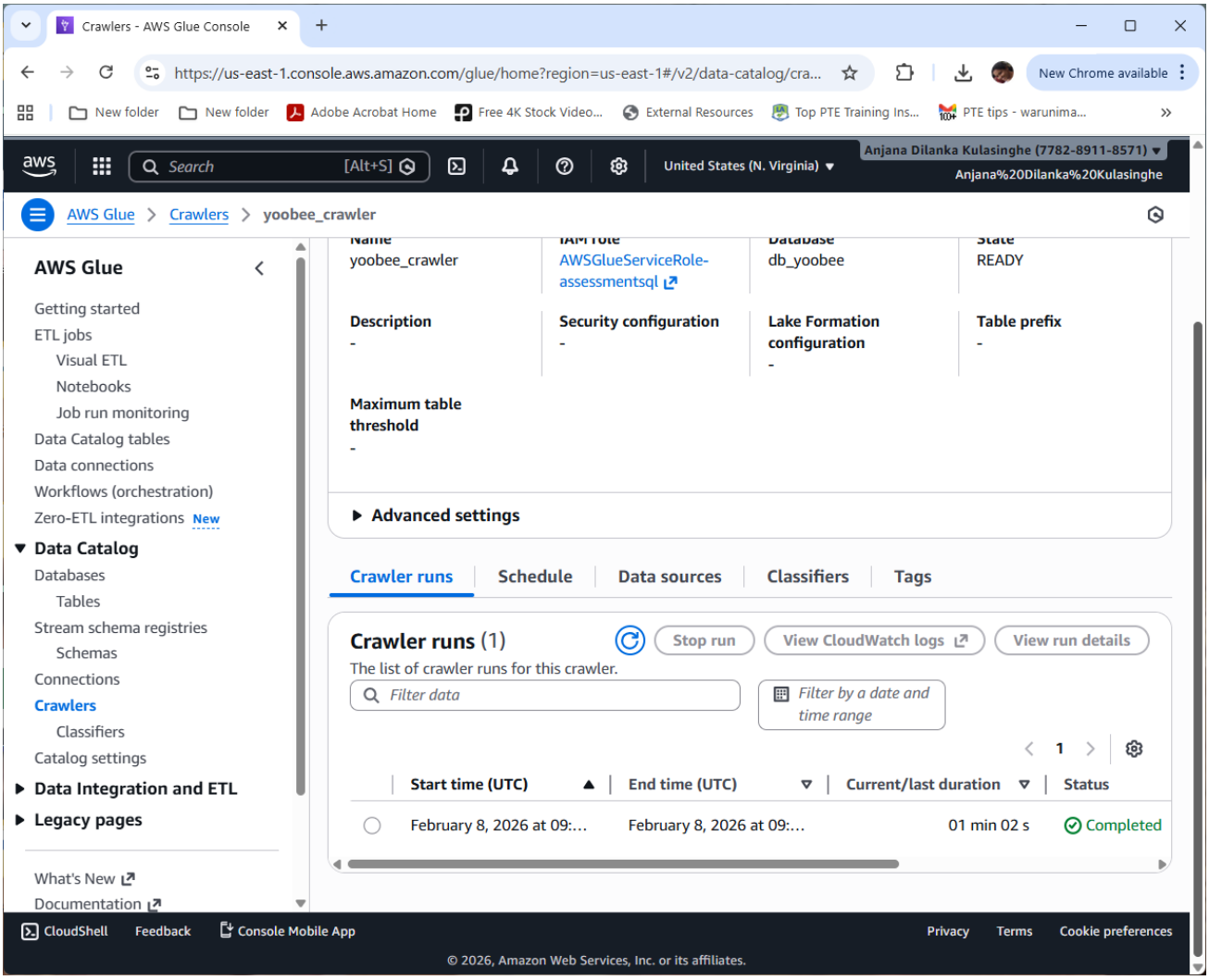
A Glue database named db_yoobee was created to manage metadata. A crawler named yoobee_crawler was configured to scan the orders/ directory inside the S3 bucket. The crawler was assigned the appropriate IAM role to allow access to S3.
After creating an additional folder (snapshot_day=2017-01-04) and uploading another dataset, the crawler was re-run. After re-running the crawler, I could see the new partitions listed in the console. I did not need to manually edit the schema, which showed how Glue simplifies metadata management.


Databases - AWS Glue Console
https://us-east-1.console.aws.amazon.com/glue/home?region=us-east-1#/v2/...
New Chrome available
New folder
New folder
Adobe Acrobat Home
Free 4K Stock Video...
External Resources
Top PTE Training Ins...
aws
Search
[Alt+S]
United States
Anjana Dilanka Kulasinghe (7782-8911-8571)
Anjana%20Dilanka%20Kulasinghe
AWS Glue
Databases
AWS Glue
Getting started
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Zero-ETL integrations New
Data Catalog
Databases
Tables
Stream schema registries
Schemas
Connections
Crawlers
Classifiers
Catalog settings
Data Integration and ETL
Legacy pages
What's New
Documentation
AWS Marketplace
Databases (1)
Last updated (UTC)
February 8, 2026 at 08:58:41
Edit
Delete
Add database
A database is a set of associated table definitions, organized into a logical group.
Filter databases
1
Name
Description
Location ...
Source ca...
Created ...
db_yoobee
-
-
778289118571
February 8, 202
ENG
US
9:58 pm
8/02/2026


AWS Glue | us-east-1
https://us-east-1.console.aws.amazon.com/glue/home?region=us-east-1#/v2/data-catalog/cra...
New Chrome available
New folder
New folder
Adobe Acrobat Home
Free 4K Stock Video...
External Resources
Top PTE Training Ins...
PTE tips - warunima...
Search
[Alt+S]
United States (N. Virginia)
Anjana Dilanka Kulasinghe (7782-8911-8571)
Anjana%20Dilanka%20Kulasinghe
AWS Glue
Crawlers
Add crawler
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Zero-ETL integrations New
Data Catalog
Databases
Tables
Stream schema registries
Schemas
Connections
Crawlers
Classifiers
Catalog settings
Data Integration and ETL
Legacy pages
What's New
Documentation
AWS Marketplace
Enable compact mode
Enable new navigation
IAM Role "AWSGlueServiceRole-assessmentsql" successfully created
Successfully created IAM Role "AWSGlueServiceRole-assessmentsql". This role trusts AWS Glue and has permissions to access your AWS Glue Crawler targets.
Step 1
Set crawler properties
Step 2
Choose data sources and classifiers
Step 3
Configure security settings
Step 4
Set output and scheduling
Step 5
Review and create
Review and create
Step 1: Set crawler properties
Edit
Set crawler properties
Name
yoobee_crawler
Description
-
Tags
-
Step 2: Choose data sources and classifiers
Edit
Data sources (1) Info
The list of data sources to be scanned by the crawler.
Type
Data source
Parameters
S3
s3://yoobee-ass...
Recrawl all
Step 3: Configure security settings
Edit
Configure security settings
IAM role
AWSGlueServiceRole-
Security configuration
-
CloudShell
Feedback
Console Mobile App
Privacy
Terms
Cookie preferences
© 2026, Amazon Web Services, Inc. or its affiliates.


Crawlers - AWS Glue Console
https://us-east-1.console.aws.amazon.com/glue/home?region=us-east-1#/v2/data-catalog/cra...
New Chrome available
New folder
New folder
Adobe Acrobat Home
Free 4K Stock Video...
External Resources
Top PTE Training Ins...
PTE tips - warunima...
aws
Search
[Alt+S]
United States (N. Virginia)
Anjana Dilanka Kulasinghe (7782-8911-8571)
Anjana%20Dilanka%20Kulasinghe
AWS Glue
Crawlers
yoobee_crawler
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Zero-ETL integrations New
Data Catalog
Databases
Tables
Stream schema registries
Schemas
Connections
Crawlers
Classifiers
Catalog settings
Data Integration and ETL
Legacy pages
What's New
Documentation
AWS Marketplace
Enable compact mode
Enable new navigation
One crawler successfully created
The following crawler is now created: "yoobee_crawler"
yoobee_crawler
Last updated (UTC)
February 8, 2026 at 09:06:48
Run crawler
Edit
Delete
Crawler properties
Name
yoobee_crawler
IAM role
AWSGlueServiceRole-assessmentsql
Database
db_yoobee
State
READY
Description
-
Security configuration
-
Lake Formation configuration
-
Table prefix
-
Maximum table threshold
-
Advanced settings
Crawler runs
Schedule
Data sources
Classifiers
Tags
Crawler runs (0)
Stop run
View CloudWatch logs
View run details
The list of crawler runs for this crawler.
Filter data
Filter by a date and time range
CloudShell
Feedback
Console Mobile App
Privacy
Terms
Cookie preferences
© 2026, Amazon Web Services, Inc. or its affiliates.


Crawlers - AWS Glue Console
https://us-east-1.console.aws.amazon.com/glue/home?region=us-east-1#/v2/data-catalog/cra...
New Chrome available
New folder
New folder
Adobe Acrobat Home
Free 4K Stock Video...
External Resources
Top PTE Training Ins...
PTE tips - warunima...
Search
[Alt+S]
United States (N. Virginia)
Anjana Dilanka Kulasinghe (7782-8911-8571)
Anjana%20Dilanka%20Kulasinghe
AWS Glue
Crawlers
yoobee_crawler
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Zero-ETL integrations New
Data Catalog
Databases
Tables
Stream schema registries
Schemas
Connections
Crawlers
Classifiers
Catalog settings
Data Integration and ETL
Legacy pages
What's New
Documentation
AWS Marketplace
Enable compact mode
Enable new navigation
Crawler successfully starting
The following crawler is now starting: "yoobee_crawler"
yoobee_crawler
Last updated (UTC)
February 8, 2026 at 09:06:48
Run crawler
Edit
Delete
Crawler properties
Name
yoobee_crawler
IAM role
AWSGlueServiceRole-assessmentsql
Database
db_yoobee
State
READY
Description
-
Security configuration
-
Lake Formation configuration
-
Table prefix
-
Maximum table threshold
-
Advanced settings
Crawler runs
Schedule
Data sources
Classifiers
Tags
Crawler runs (1)
Stop run
View CloudWatch logs
View run details
The list of crawler runs for this crawler.
Filter data
Filter by a date and time range
CloudShell
Feedback
Console Mobile App
Privacy
Terms
Cookie preferences
© 2026, Amazon Web Services, Inc. or its affiliates.


Crawlers - AWS Glue Console
https://us-east-1.console.aws.amazon.com/glue/home?region=us-east-1#/v2/data-catalog/cra...
New Chrome available
New folder
New folder
Adobe Acrobat Home
Free 4K Stock Video...
External Resources
Top PTE Training Ins...
PTE tips - warunima...
Search
[Alt+S]
United States (N. Virginia)
Anjana Dilanka Kulasinghe (7782-8911-8571)
Anjana%20Dilanka%20Kulasinghe
AWS Glue
Crawlers
yoobee_crawler
AWS Glue
Getting started
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Zero-ETL integrations New
Data Catalog
Databases
Tables
Stream schema registries
Schemas
Connections
Crawlers
Classifiers
Catalog settings
Data Integration and ETL
Legacy pages
What's New
Documentation
yoobee_crawler
AWSGlueServiceRole-assessmentsql
db_yoobee
READY
Description
-
Security configuration
-
Lake Formation configuration
-
Table prefix
-
Maximum table threshold
-
Advanced settings
Crawler runs
Schedule
Data sources
Classifiers
Tags
Crawler runs (1)
Stop run
View CloudWatch logs
View run details
The list of crawler runs for this crawler.
Filter data
Filter by a date and time range
Start time (UTC)
End time (UTC)
Current/last duration
Status
February 8, 2026 at 09:...
February 8, 2026 at 09:...
01 min 02 s
Completed
CloudShell
Feedback
Console Mobile App
Privacy
Terms
Cookie preferences
© 2026, Amazon Web Services, Inc. or its affiliates.

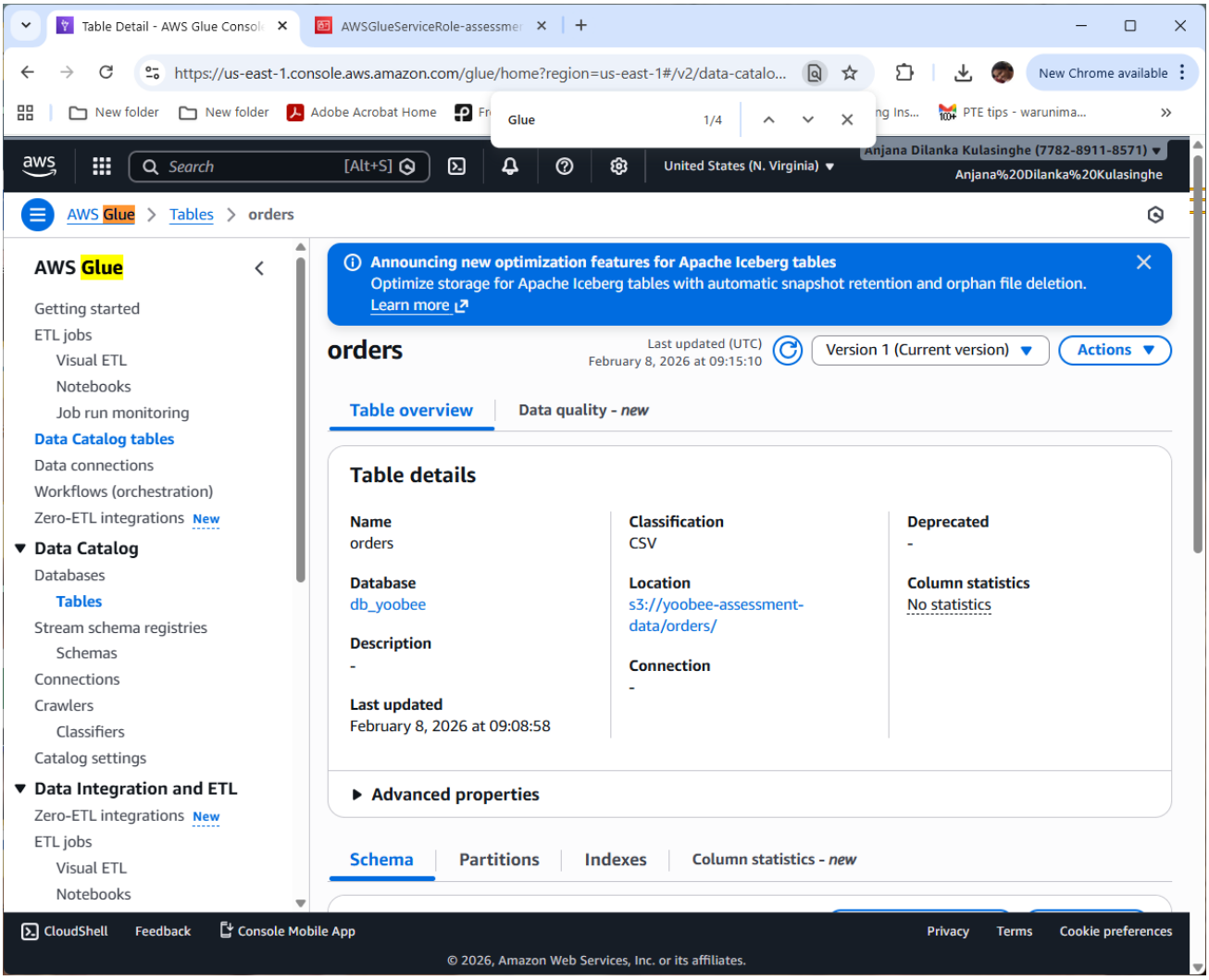


Table Detail - AWS Glue Consol
AWSGlueServiceRole-assessmer
https://us-east-1.console.aws.amazon.com/glue/home?region=us-east-1#/v2/data-catalo...
New Chrome available
New folder
New folder
Adobe Acrobat Home
Glue
1/4
PTE tips - warunima...
aws
Search
[Alt+S]
United States (N. Virginia)
Anjana Dilanka Kulasinghe (7782-8911-8571)
Anjana%20Dilanka%20Kulasinghe
AWS Glue > Tables > orders
AWS Glue
Getting started
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Zero-ETL integrations New
Data Catalog
Databases
Tables
Stream schema registries
Schemas
Connections
Crawlers
Classifiers
Catalog settings
Data Integration and ETL
Zero-ETL integrations New
ETL jobs
Visual ETL
Notebooks
Announcing new optimization features for Apache Iceberg tables
Optimize storage for Apache Iceberg tables with automatic snapshot retention and orphan file deletion.
Learn more
orders
Last updated (UTC)
February 8, 2026 at 09:15:10
Version 1 (Current version)
Actions
Table overview
Data quality - new
Table details
Name
orders
Database
db_yoobee
Description
-
Last updated
February 8, 2026 at 09:08:58
Classification
CSV
Location
s3://yoobee-assessment-data/orders/
Connection
-
Deprecated
-
Column statistics
No statistics
Advanced properties
Schema
Partitions
Indexes
Column statistics - new
CloudShell
Feedback
Console Mobile App
Privacy
Terms
Cookie preferences
© 2026, Amazon Web Services, Inc. or its affiliates.

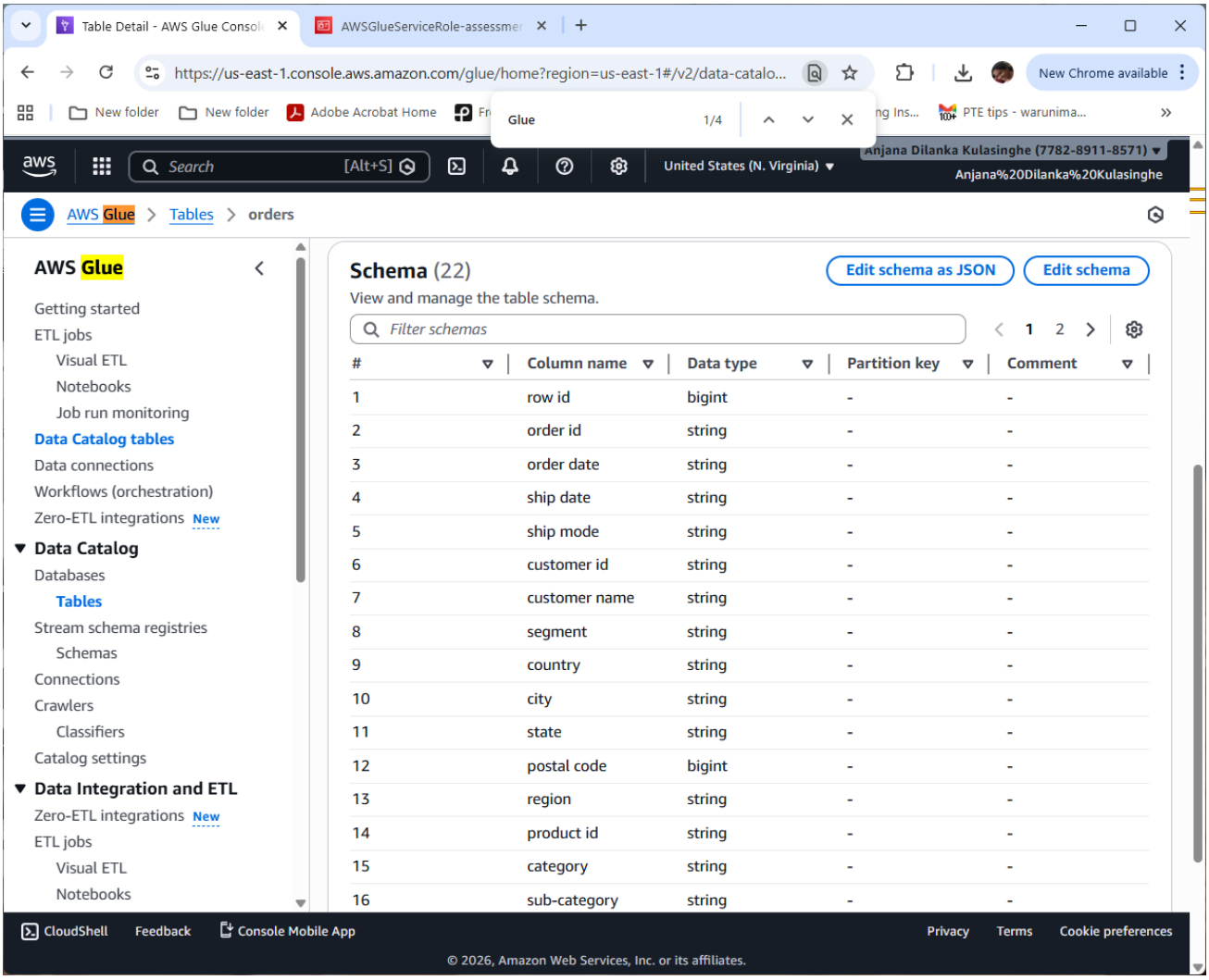


Table Detail - AWS Glue Consol
AWSGlueServiceRoleAssessmer
https://us-east-1.console.aws.amazon.com/glue/home?region=us-east-1#/v2/data-catalo...
New Chrome available
New folder
New folder
Adobe Acrobat Home
Glue
1/4
PTE tips - warunima...
aws
Search
[Alt+S]
United States (N. Virginia)
Anjana Dilanka Kulasinghe (7782-8911-8571)
Anjana%20Dilanka%20Kulasinghe
AWS Glue
Tables
orders
AWS Glue
Getting started
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Zero-ETL integrations New
Data Catalog
Databases
Tables
Stream schema registries
Schemas
Connections
Crawlers
Classifiers
Catalog settings
Data Integration and ETL
Zero-ETL integrations New
ETL jobs
Visual ETL
Notebooks
Schema (22)
View and manage the table schema.
Edit schema as JSON
Edit schema
Filter schemas
1
2
| Column name | Data type | Partition key | Comment
1 | row id | bigint | - | -
2 | order id | string | - | -
3 | order date | string | - | -
4 | ship date | string | - | -
5 | ship mode | string | - | -
6 | customer id | string | - | -
7 | customer name | string | - | -
8 | segment | string | - | -
9 | country | string | - | -
10 | city | string | - | -
11 | state | string | - | -
12 | postal code | bigint | - | -
13 | region | string | - | -
14 | product id | string | - | -
15 | category | string | - | -
16 | sub-category | string | - | -
CloudShell
Feedback
Console Mobile App
Privacy
Terms
Cookie preferences
© 2026, Amazon Web Services, Inc. or its affiliates.

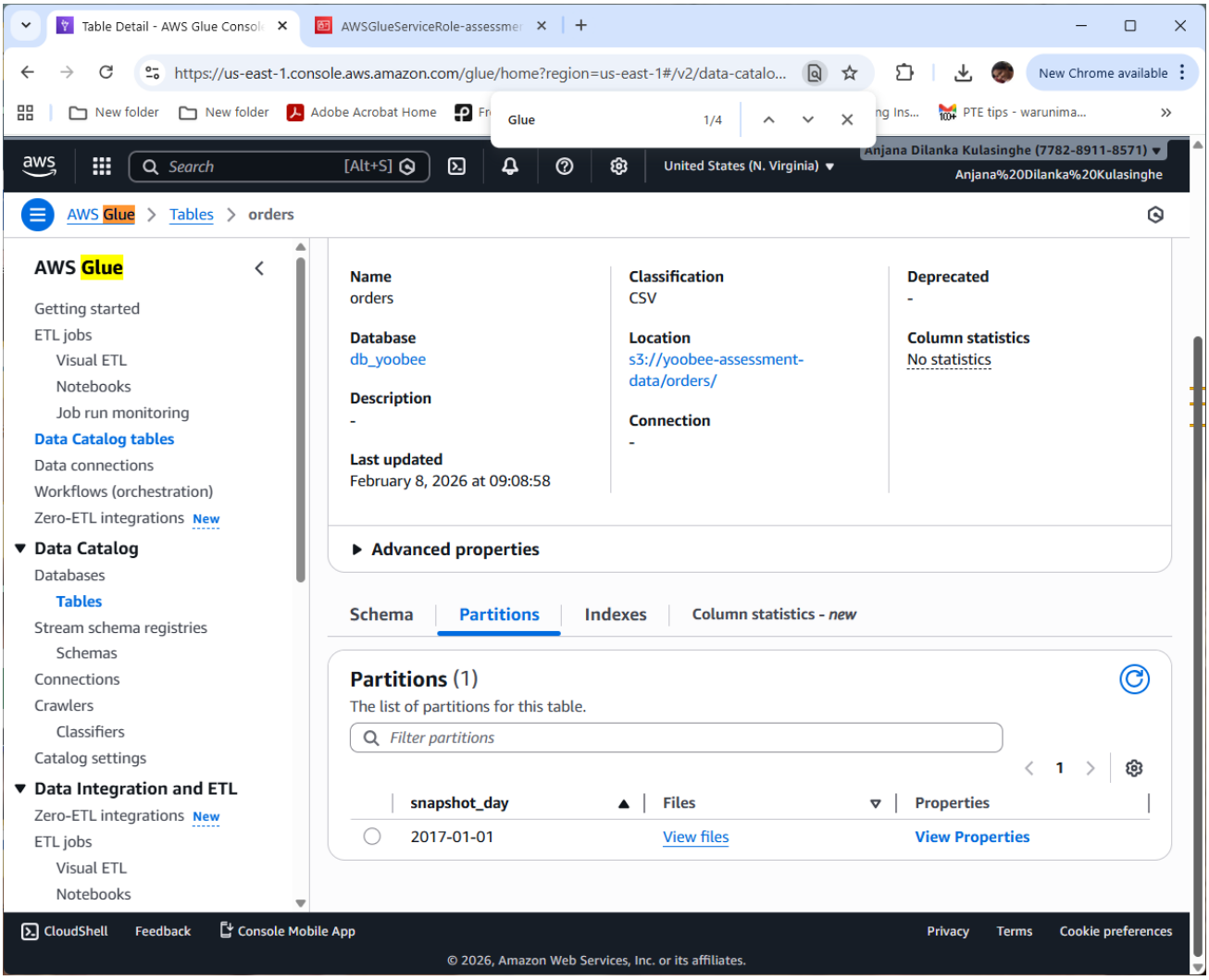
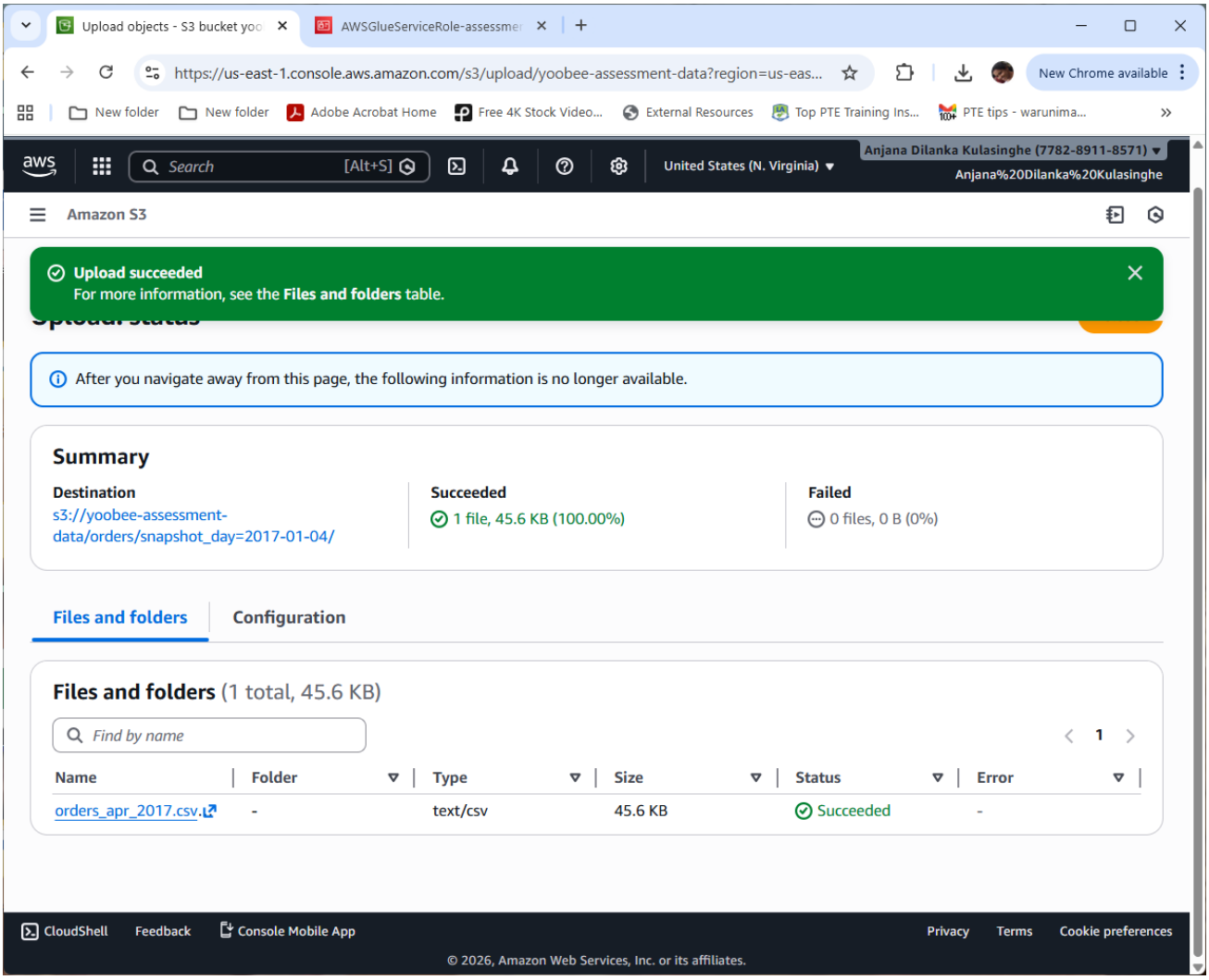
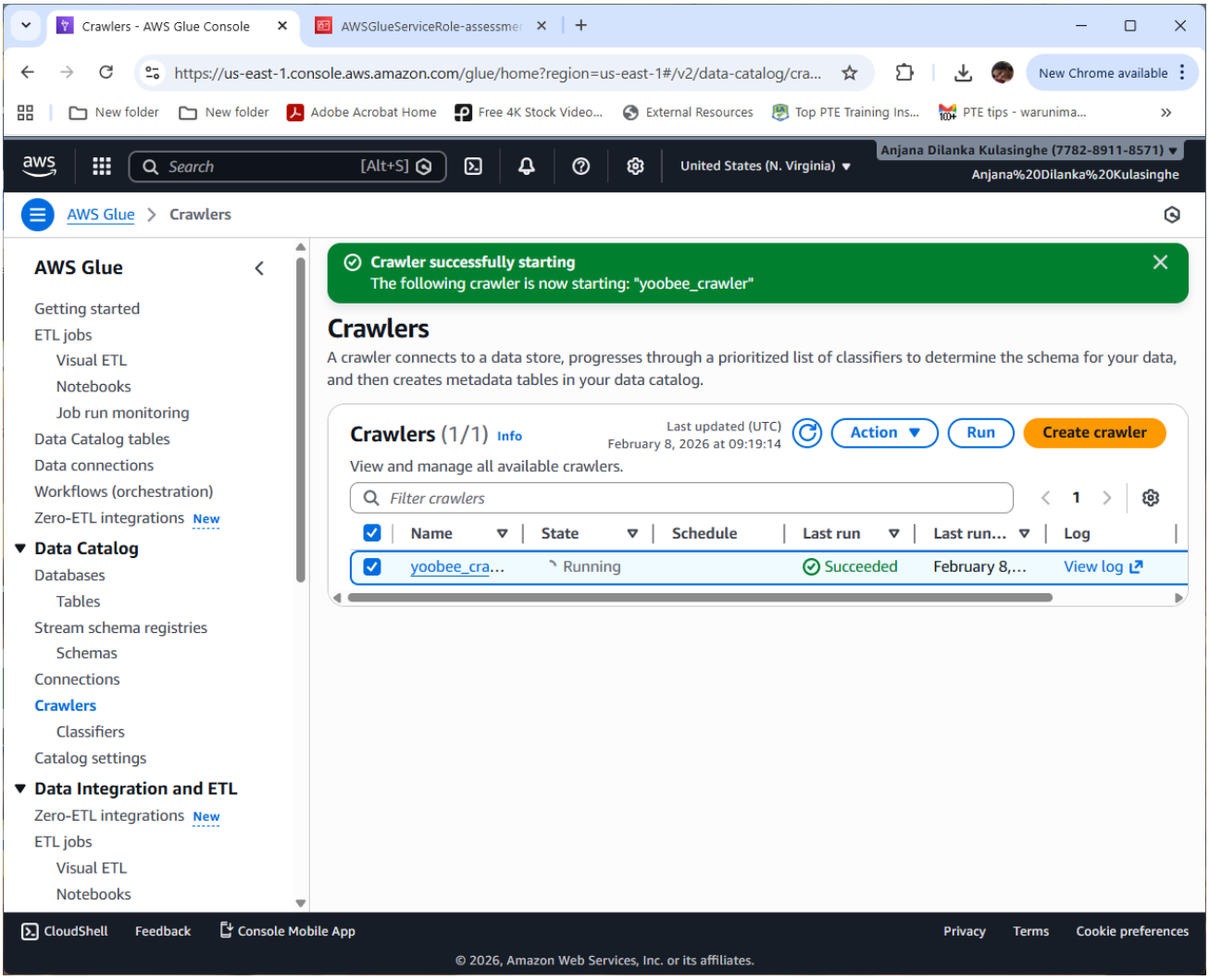
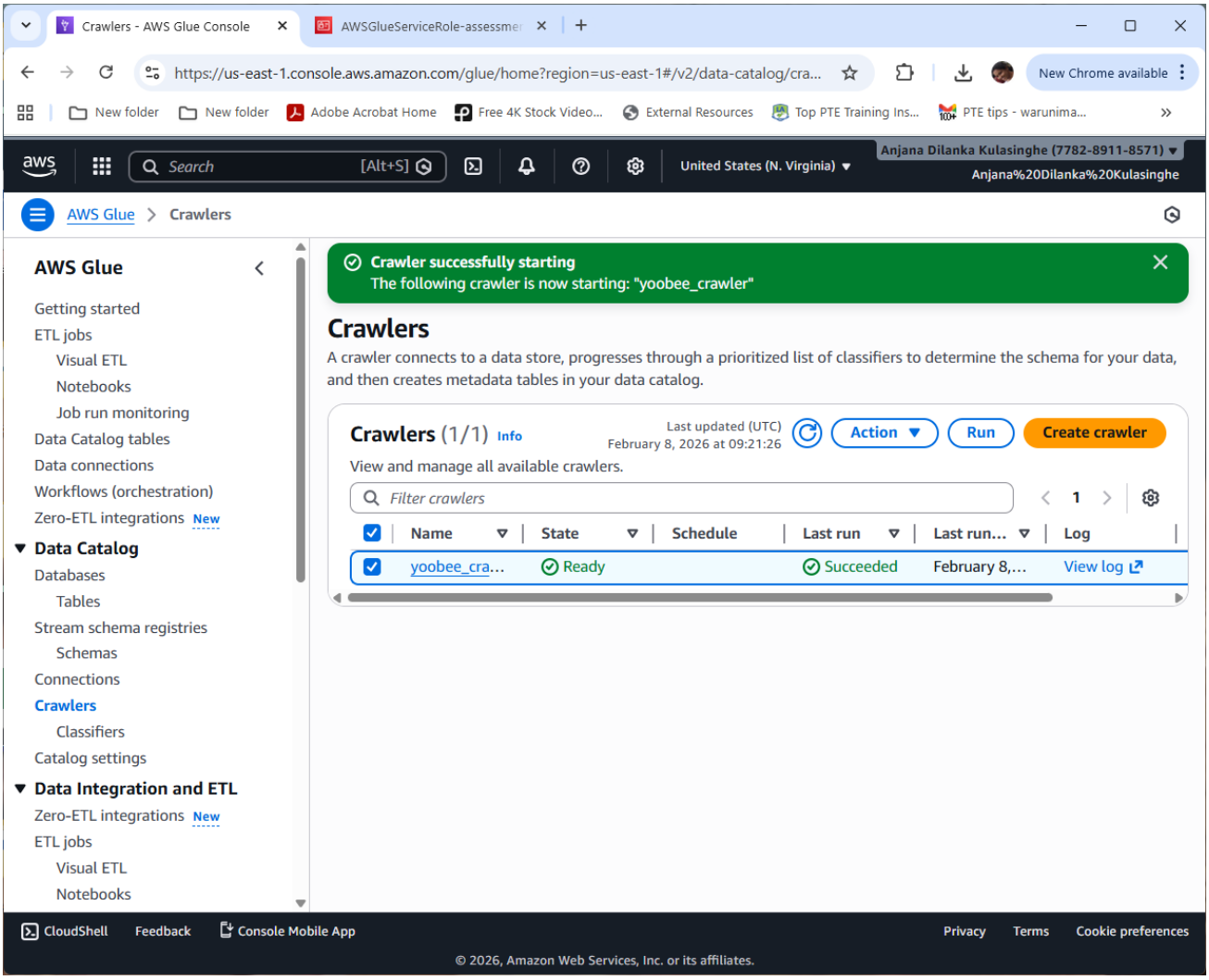
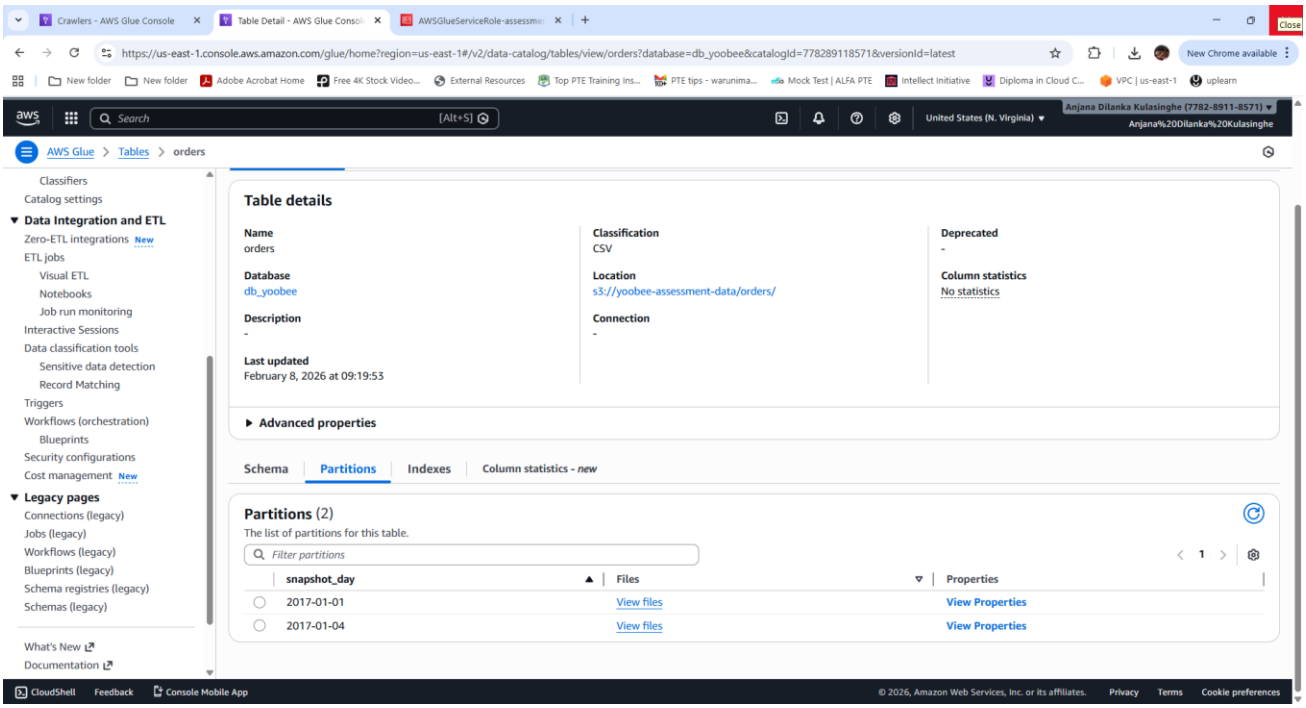


Table Detail - AWS Glue Consol
AWSGlueServiceRoleAssessmer
https://us-east-1.console.aws.amazon.com/glue/home?region=us-east-1#/v2/data-catalo...
New Chrome available
New folder
New folder
Adobe Acrobat Home
Glue
1/4
PTE tips - warunima...
aws
Search
[Alt+S]
United States (N. Virginia)
Anjana Dilanka Kulasinghe (7782-8911-8571)
Anjana%20Dilanka%20Kulasinghe
AWS Glue
Tables
orders
AWS Glue
Getting started
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Zero-ETL integrations New
Data Catalog
Databases
Tables
Stream schema registries
Schemas
Connections
Crawlers
Classifiers
Catalog settings
Data Integration and ETL
Zero-ETL integrations New
ETL jobs
Visual ETL
Notebooks
Name
orders
Database
db_yoobee
Description
-
Last updated
February 8, 2026 at 09:08:58
Classification
CSV
Location
s3://yoobee-assessment-data/orders/
Connection
-
Deprecated
-
Column statistics
No statistics
Advanced properties
Schema
Partitions
Indexes
Column statistics - new
Partitions (1)
The list of partitions for this table.
Filter partitions
snapshot_day
Files
Properties
2017-01-01
View files
View Properties
CloudShell
Feedback
Console Mobile App
Privacy
Terms
Cookie preferences
© 2026, Amazon Web Services, Inc. or its affiliates.


Upload objects - S3 bucket yoo
AWSGlueServiceRole-assessmer
https://us-east-1.console.aws.amazon.com/s3/upload/yoobee-assessment-data?region=us-eas...
New Chrome available
New folder
New folder
Adobe Acrobat Home
Free 4K Stock Video...
External Resources
Top PTE Training Ins...
PTE tips - warunima...
aws
Search
[Alt+S]
United States (N. Virginia)
Anjana Dilanka Kulasinghe (7782-8911-8571)
Anjana%20Dilanka%20Kulasinghe
Amazon S3
Upload succeeded
For more information, see the Files and folders table.
After you navigate away from this page, the following information is no longer available.
Summary
Destination
s3://yoobee-assessment-data/orders/snapshot_day=2017-01-04/
Succeeded
1 file, 45.6 KB (100.00%)
Failed
0 files, 0 B (0%)
Files and folders
Configuration
Files and folders (1 total, 45.6 KB)
Find by name
1
Name
Folder
Type
Size
Status
Error
orders_apr_2017.csv
-
text/csv
45.6 KB
Succeeded
-
CloudShell
Feedback
Console Mobile App
Privacy
Terms
Cookie preferences
© 2026, Amazon Web Services, Inc. or its affiliates.


Crawlers - AWS Glue Console
AWSGlueServiceRoleAssessmer
https://us-east-1.console.aws.amazon.com/glue/home?region=us-east-1#/v2/data-catalog/cra...
New Chrome available
New folder
New folder
Adobe Acrobat Home
Free 4K Stock Video...
External Resources
Top PTE Training Ins...
PTE tips - warunima...
aws
Search
[Alt+S]
United States (N. Virginia)
Anjana Dilanka Kulasinghe (7782-8911-8571)
Anjana%20Dilanka%20Kulasinghe
AWS Glue
Crawlers
AWS Glue
Getting started
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Zero-ETL integrations New
Data Catalog
Databases
Tables
Stream schema registries
Schemas
Connections
Crawlers
Classifiers
Catalog settings
Data Integration and ETL
Zero-ETL integrations New
ETL jobs
Visual ETL
Notebooks
Crawler successfully starting
The following crawler is now starting: "yoobee_crawler"
Crawlers
A crawler connects to a data store, progresses through a prioritized list of classifiers to determine the schema for your data, and then creates metadata tables in your data catalog.
Crawlers (1/1) Info
Last updated (UTC)
February 8, 2026 at 09:19:14
Action
Run
Create crawler
View and manage all available crawlers.
Filter crawlers
1
Name
State
Schedule
Last run
Last run...
Log
yoobee_cra...
Running
Succeeded
February 8,...
View log
CloudShell
Feedback
Console Mobile App
Privacy
Terms
Cookie preferences
© 2026, Amazon Web Services, Inc. or its affiliates.


Crawlers - AWS Glue Console
AWSGlueServiceRoleAssessmer
https://us-east-1.console.aws.amazon.com/glue/home?region=us-east-1#/v2/data-catalog/cra...
New Chrome available
New folder
New folder
Adobe Acrobat Home
Free 4K Stock Video...
External Resources
Top PTE Training Ins...
PTE tips - warunima...
Search
[Alt+S]
United States (N. Virginia)
Anjana Dilanka Kulasinghe (7782-8911-8571)
Anjana%20Dilanka%20Kulasinghe
AWS Glue
Crawlers
AWS Glue
Getting started
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Zero-ETL integrations New
Data Catalog
Databases
Tables
Stream schema registries
Schemas
Connections
Crawlers
Classifiers
Catalog settings
Data Integration and ETL
Zero-ETL integrations New
ETL jobs
Visual ETL
Notebooks
Crawler successfully starting
The following crawler is now starting: "yoobee_crawler"
Crawlers
A crawler connects to a data store, progresses through a prioritized list of classifiers to determine the schema for your data, and then creates metadata tables in your data catalog.
Crawlers (1/1) Info
Last updated (UTC)
February 8, 2026 at 09:21:26
Action
Run
Create crawler
View and manage all available crawlers.
Filter crawlers
Name
State
Schedule
Last run
Last run...
Log
yoobee_cra...
Ready
Succeeded
February 8,...
View log
CloudShell
Feedback
Console Mobile App
Privacy
Terms
Cookie preferences
© 2026, Amazon Web Services, Inc. or its affiliates.

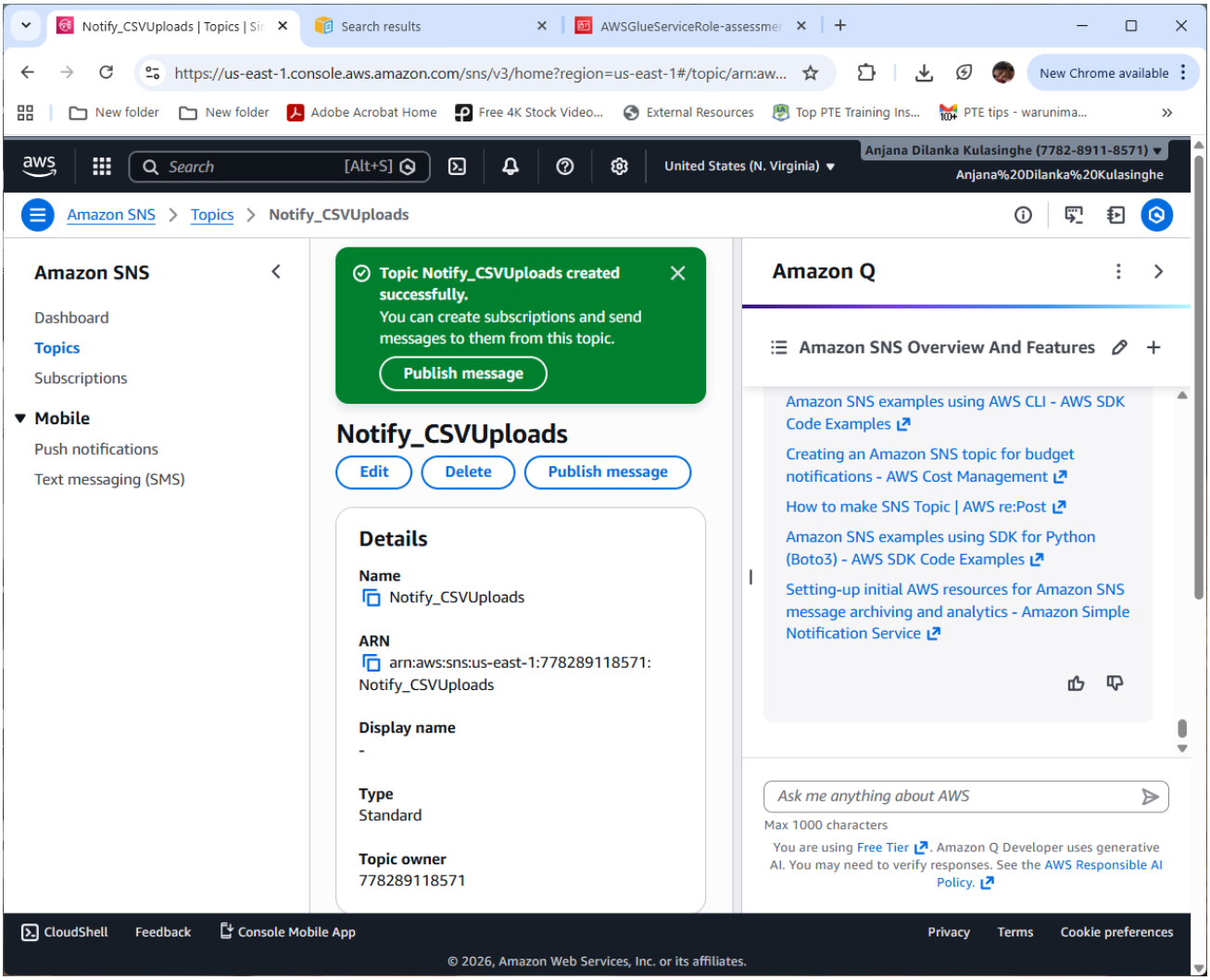

AWS Glue
Tables
orders
Table details
Name
orders
Classification
CSV
Deprecated
-
Database
db_yoobee
Location
s3://yoobee-assessment-data/orders/
Column statistics
No statistics
Description
-
Connection
-
Last updated
February 8, 2026 at 09:19:53
Advanced properties
Schema
Partitions
Indexes
Column statistics - new
Partitions (2)
The list of partitions for this table.
Filter partitions
snapshot_day
Files
Properties
2017-01-01
View files
View Properties
2017-01-04
View files
View Properties

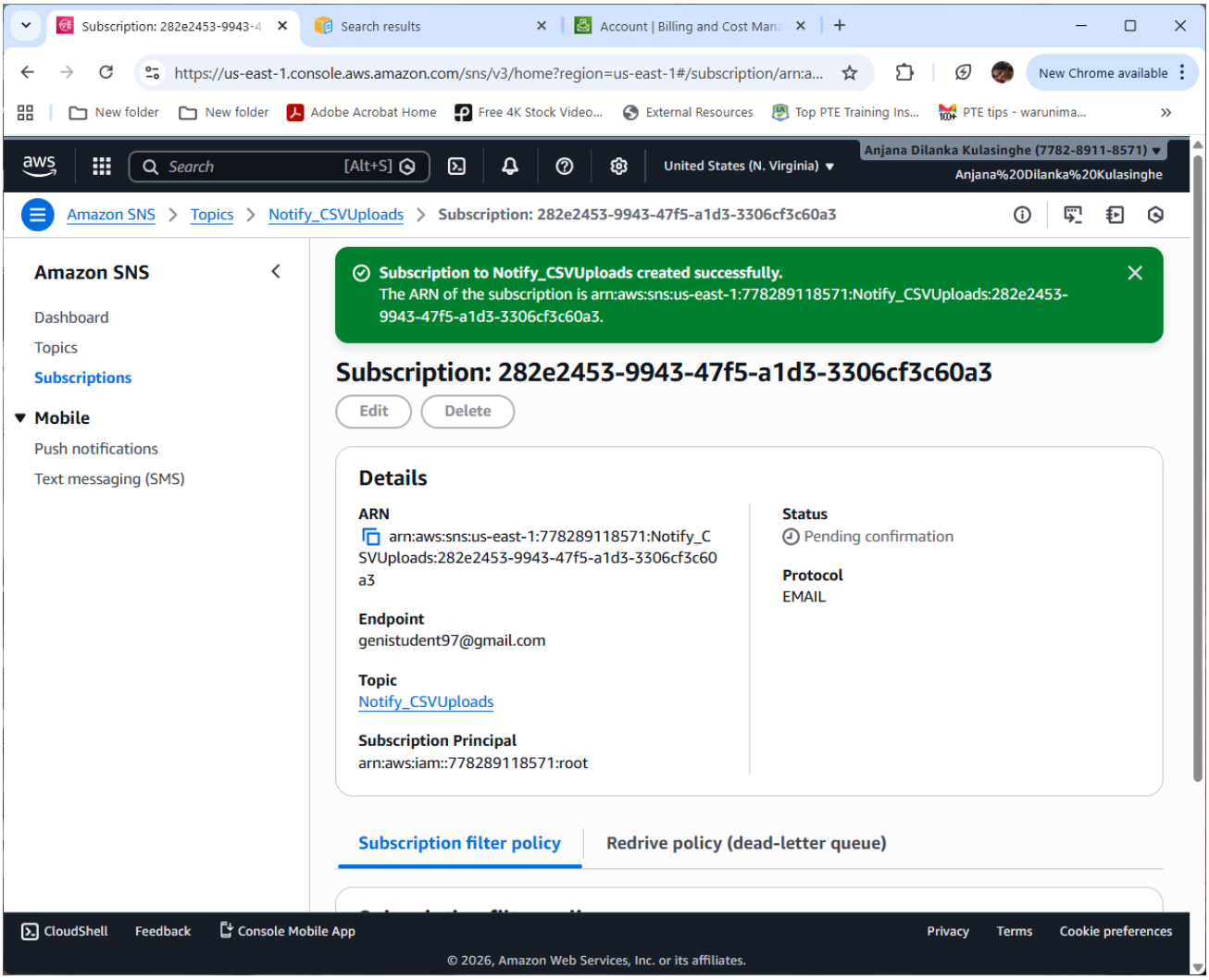
Activities 1.3: Real-Time Notification Using AWS Lambda and SNS

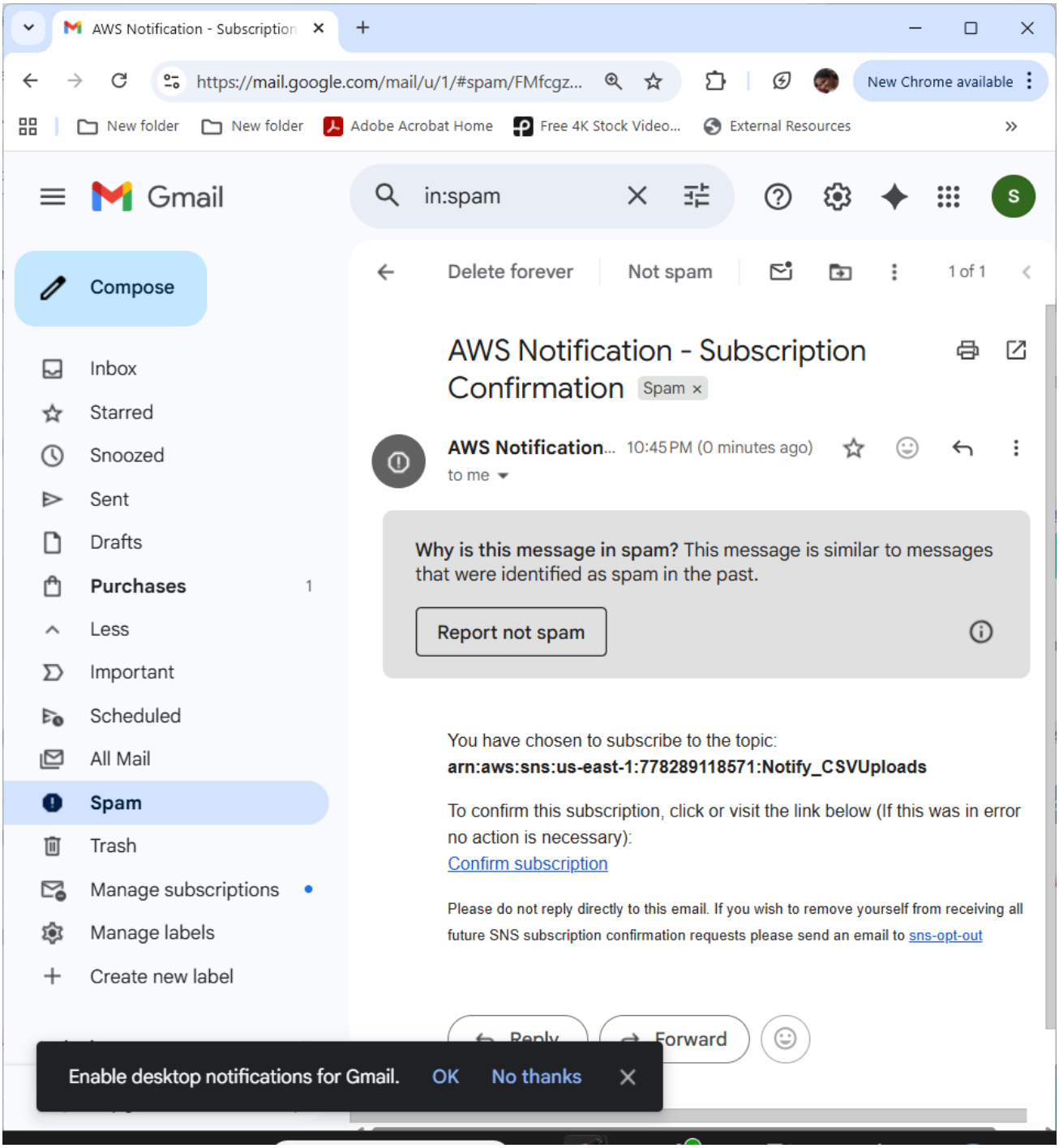
To implement real-time monitoring, an SNS topic named Notify CSVUploads was created. A Lambda function was then deployed and configured with an S3 ObjectCreated trigger. The function checks whether the uploaded file has a .csv extension before publishing a structured notification message to the SNS topic.

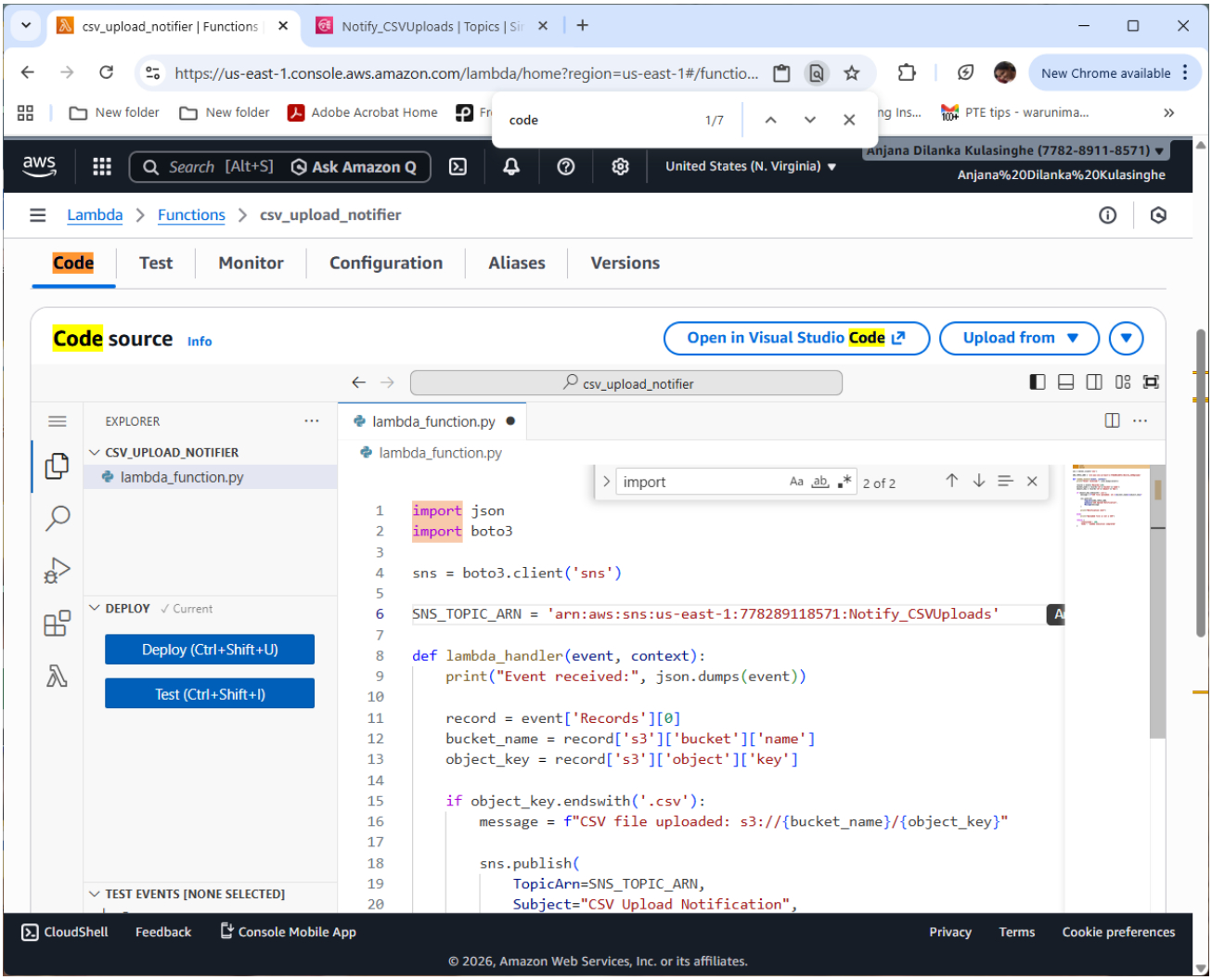
This setup demonstrates event-driven automation. Instead of manually checking for new files, the system automatically reacts whenever new data is uploaded.

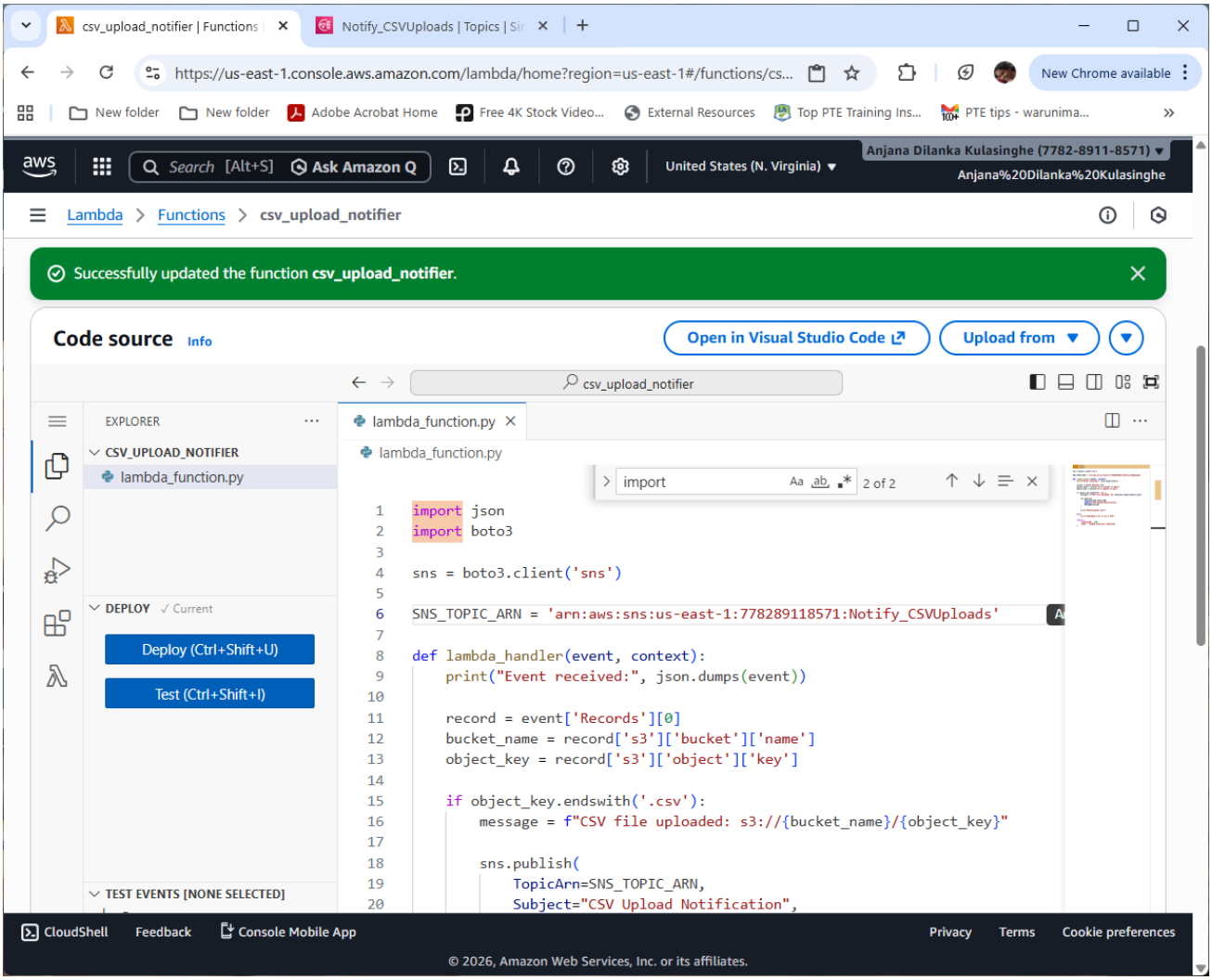
I tested this by uploading a new CSV file, and the notification was triggered immediately.

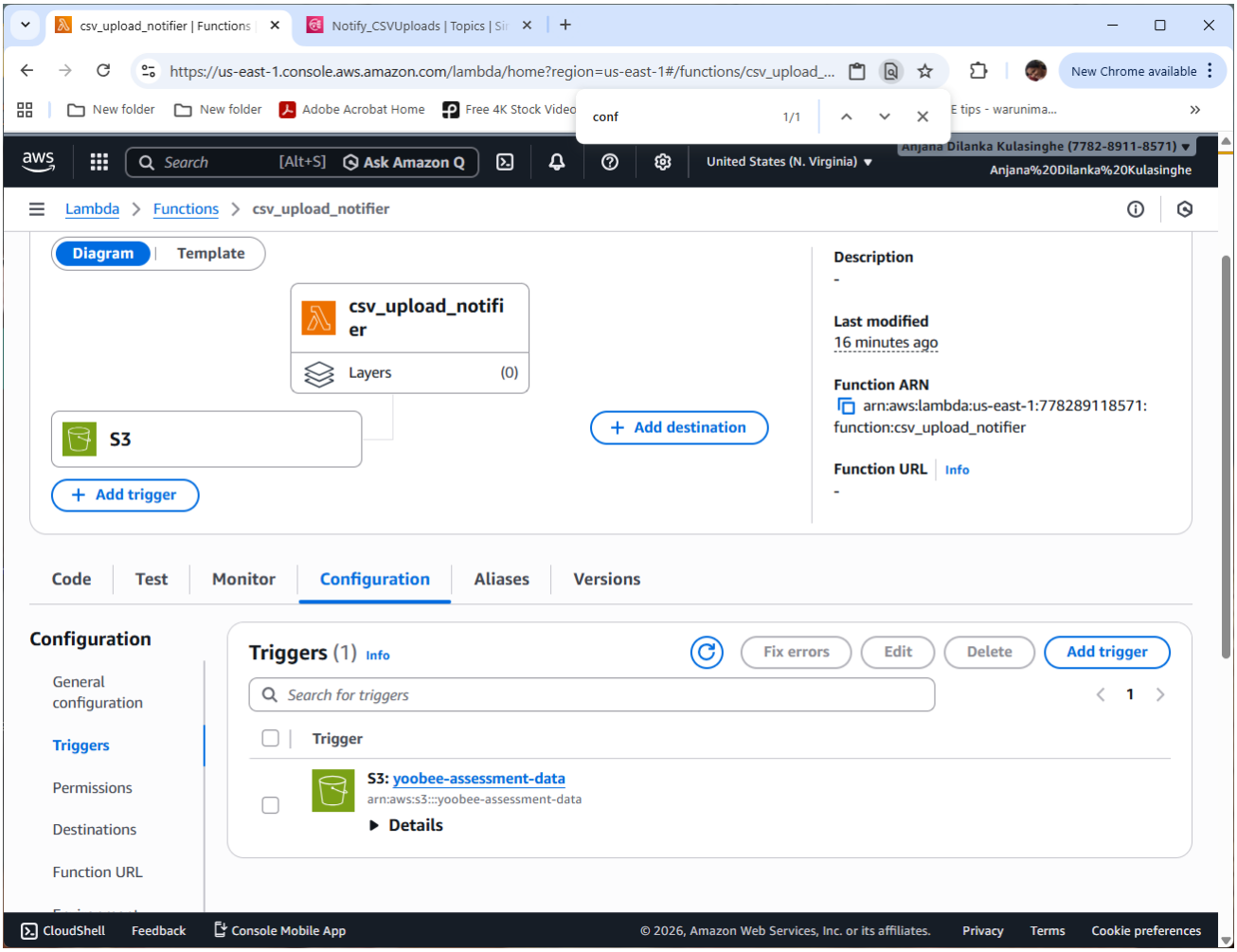


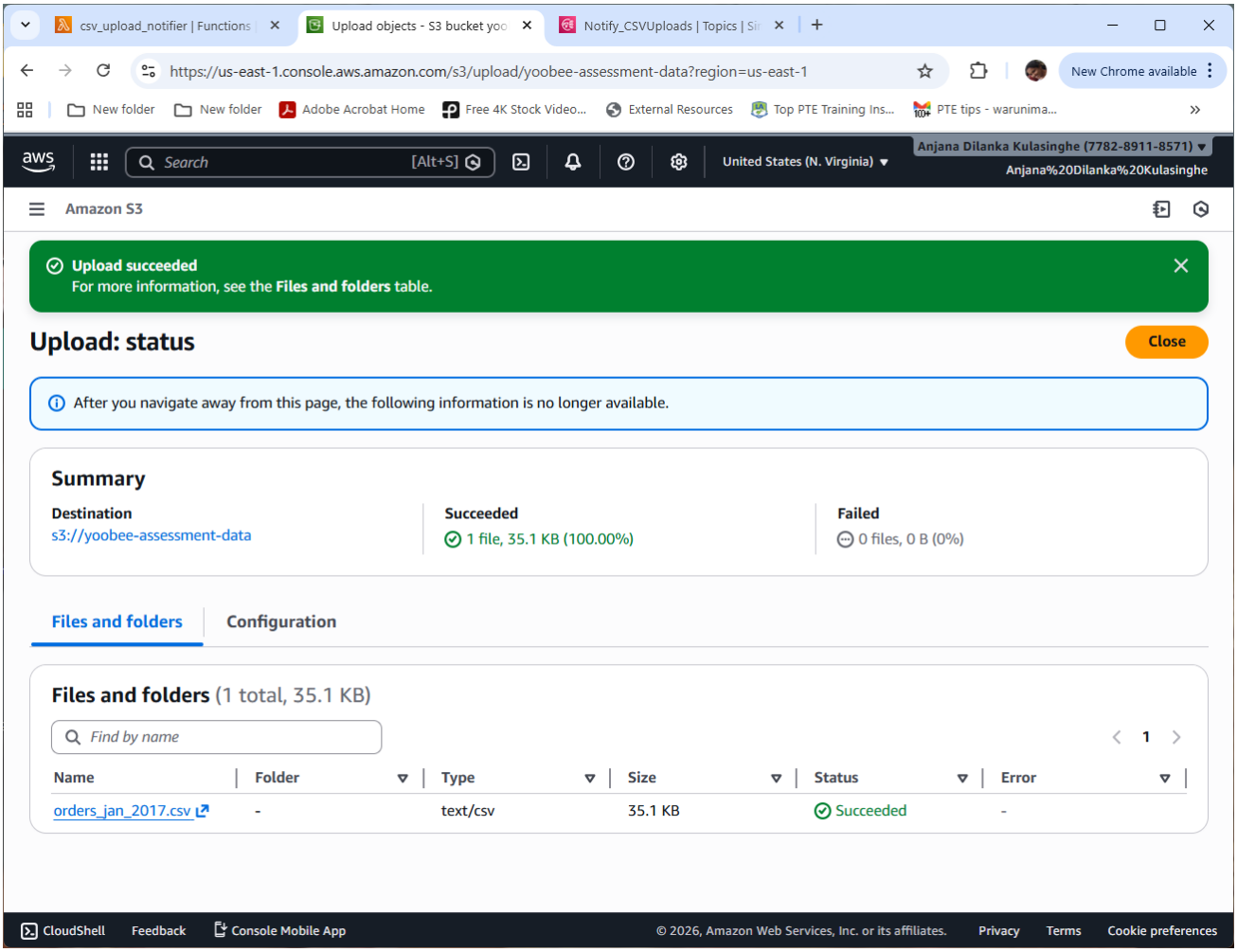

Subscription: 282e2453-9943-4
Search results
Account | Billing and Cost Mana
https://us-east-1.console.aws.amazon.com/sns/v3/home?region=us-east-1#/subscription/arn:a...
New Chrome available
New folder
New folder
Adobe Acrobat Home
Free 4K Stock Video...
External Resources
Top PTE Training Ins...
PTE tips - warunima...
aws
Search
[Alt+S]
United States (N. Virginia)
Anjana Dilanka Kulasinghe (7782-8911-8571)
Anjana%20Dilanka%20Kulasinghe
Amazon SNS
Topics
Notify_CSVUploads
Subscription: 282e2453-9943-47f5-a1d3-3306cf3c60a3
Amazon SNS
Dashboard
Topics
Subscriptions
Mobile
Push notifications
Text messaging (SMS)
Subscription to Notify_CSVUploads created successfully.
The ARN of the subscription is arn:aws:sns:us-east-1:778289118571:Notify_CSVUploads:282e2453-9943-47f5-a1d3-3306cf3c60a3.
Subscription: 282e2453-9943-47f5-a1d3-3306cf3c60a3
Edit
Delete
Details
ARN
arn:aws:sns:us-east-1:778289118571:Notify_CSVUploads:282e2453-9943-47f5-a1d3-3306cf3c60a3
Endpoint
genistudent97@gmail.com
Topic
Notify_CSVUploads
Subscription Principal
arn:aws:iam::778289118571:root
Status
Pending confirmation
Protocol
EMAIL
Subscription filter policy
Redrive policy (dead-letter queue)
CloudShell
Feedback
Console Mobile App
Privacy
Terms
Cookie preferences
© 2026, Amazon Web Services, Inc. or its affiliates.

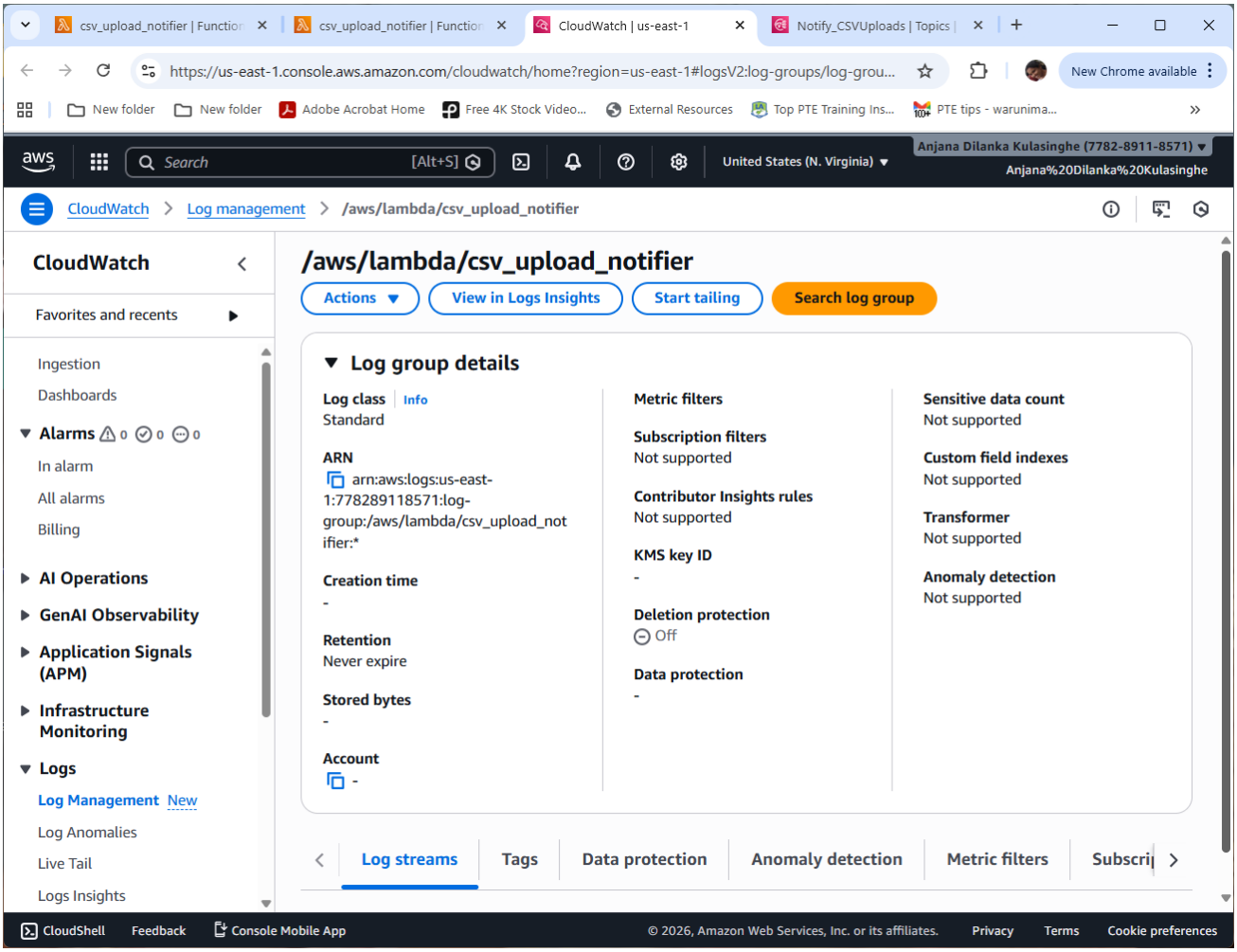

AWS Notification - Subscription
https://mail.google.com/mail/u/1/#spam/FMfcgz...
New Chrome available
New folder
New folder
Adobe Acrobat Home
Free 4K Stock Video...
External Resources
Gmail
in:spam
Compose
Inbox
Starred
Snoozed
Sent
Drafts
Purchases 1
Less
Important
Scheduled
All Mail
Spam
Trash
Manage subscriptions
Manage labels
Create new label
Delete forever
Not spam
1 of 1
AWS Notification - Subscription Confirmation
Spam
AWS Notification... 10:45 PM (0 minutes ago)
to me
Why is this message in spam? This message is similar to messages that were identified as spam in the past.
Report not spam
You have chosen to subscribe to the topic:
arn:aws:sns:us-east-1:778289118571:Notify_CSVUploads
To confirm this subscription, click or visit the link below (If this was in error no action is necessary):
Confirm subscription
Please do not reply directly to this email. If you wish to remove yourself from receiving all future SNS subscription confirmation requests please send an email to sns-opt-out
Reply
Forward
Enable desktop notifications for Gmail.
OK
No thanks


csv_upload_notifier | Functions
Notify_CSVUploads | Topics
https://us-east-1.console.aws.amazon.com/lambda/home?region=us-east-1#/functio...
New Chrome available
code 1/7
Search [Alt+S]
Ask Amazon Q
United States (N. Virginia)
Anjana Dilanka Kulasinghe (7782-8911-8571)
Anjana%20Dilanka%20Kulasinghe
Lambda > Functions > csv_upload_notifier
Code
Test
Monitor
Configuration
Aliases
Versions
Code source Info
Open in Visual Studio Code
Upload from
csv_upload_notifier
EXPLORER
CSV_UPLOAD_NOTIFIER
lambda_function.py
DEPLOY
Current
Deploy (Ctrl+Shift+U)
Test (Ctrl+Shift+I)
TEST EVENTS [NONE SELECTED]
import 2 of 2
import json
import boto3
sns = boto3.client('sns')
SNS_TOPIC_ARN = 'arn:aws:sns:us-east-1:778289118571:Notify_CSVUploads'
def lambda_handler(event, context):
print("Event received:", json.dumps(event))
record = event['Records'][0]
bucket_name = record['s3']['bucket']['name']
object_key = record['s3']['object']['key']
if object_key.endswith('.csv'):
message = f"CSV file uploaded: s3://{bucket_name}/{object_key}"
sns.publish(
TopicArn=SNS_TOPIC_ARN,
Subject="CSV Upload Notification",
CloudShell
Feedback
Console Mobile App
Privacy
Terms
Cookie preferences
© 2026, Amazon Web Services, Inc. or its affiliates.


csv_upload_notifier | Functions
Notify_CSVUploads | Topics | Sir
https://us-east-1.console.aws.amazon.com/lambda/home?region=us-east-1#/functions/cs...
New Chrome available
New folder
New folder
Adobe Acrobat Home
Free 4K Stock Video...
External Resources
Top PTE Training Ins...
PTE tips - warunima...
aws
Search [Alt+S]
Ask Amazon Q
United States (N. Virginia)
Anjana Dilanka Kulasinghe (7782-8911-8571)
Anjana%20Dilanka%20Kulasinghe
Lambda > Functions > csv_upload_notifier
Successfully updated the function csv_upload_notifier.
Code source Info
Open in Visual Studio Code
Upload from
csv_upload_notifier
EXPLORER
CSV_UPLOAD_NOTIFIER
lambda_function.py
DEPLOY ✓ Current
Deploy (Ctrl+Shift+U)
Test (Ctrl+Shift+I)
TEST EVENTS [NONE SELECTED]
lambda_function.py
import
2 of 2
1 import json
2 import boto3
3
4 sns = boto3.client('sns')
5
6 SNS_TOPIC_ARN = 'arn:aws:sns:us-east-1:778289118571:Notify_CSVUploads'
7
8 def lambda_handler(event, context):
9 print("Event received:", json.dumps(event))
10
11 record = event['Records'][0]
12 bucket_name = record['s3']['bucket']['name']
13 object_key = record['s3']['object']['key']
14
15 if object_key.endswith('.csv'):
16 message = f"CSV file uploaded: s3://{bucket_name}/{object_key}"
17
18 sns.publish(
19 TopicArn=SNS_TOPIC_ARN,
20 Subject="CSV Upload Notification",
CloudShell
Feedback
Console Mobile App
Privacy
Terms
Cookie preferences
© 2026, Amazon Web Services, Inc. or its affiliates.


csv_upload_notifier | Functions
Notify_CSVUploads | Topics
https://us-east-1.console.aws.amazon.com/lambda/home?region=us-east-1#/functions/csv_upload_...
New Chrome available
conf
1/1
Search
[Alt+S]
Ask Amazon Q
United States (N. Virginia)
Anjana Dilanka Kulasinghe (7782-8911-8571)
Anjana%20Dilanka%20Kulasinghe
Lambda > Functions > csv_upload_notifier
Diagram
Template
csv_upload_notifier
Layers
(0)
S3
Add trigger
Add destination
Description
-
Last modified
16 minutes ago
Function ARN
arn:aws:lambda:us-east-1:778289118571:function:csv_upload_notifier
Function URL
Info
-
Code
Test
Monitor
Configuration
Aliases
Versions
Configuration
General configuration
Triggers
Permissions
Destinations
Function URL
Triggers (1) Info
Fix errors
Edit
Delete
Add trigger
Search for triggers
Trigger
S3: yoobee-assessment-data
arn:aws:s3:::yoobee-assessment-data
Details
CloudShell
Feedback
Console Mobile App
© 2026, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences


csv_upload_notifier | Functions
Upload objects - S3 bucket yoo
Notify_CSVUploads | Topics
https://us-east-1.console.aws.amazon.com/s3/upload/yoobee-assessment-data?region=us-east-1
New Chrome available
New folder
New folder
Adobe Acrobat Home
Free 4K Stock Video...
External Resources
Top PTE Training Ins...
PTE tips - warunima...
aws
Search
[Alt+S]
United States (N. Virginia)
Anjana Dilanka Kulasinghe (7782-8911-8571)
Anjana%20Dilanka%20Kulasinghe
Amazon S3
Upload succeeded
For more information, see the Files and folders table.
Upload: status
Close
After you navigate away from this page, the following information is no longer available.
Summary
Destination
s3://yoobee-assessment-data
Succeeded
1 file, 35.1 KB (100.00%)
Failed
0 files, 0 B (0%)
Files and folders
Configuration
Files and folders (1 total, 35.1 KB)
Find by name
1
Name
Folder
Type
Size
Status
Error
orders_jan_2017.csv
-
text/csv
35.1 KB
Succeeded
-
CloudShell
Feedback
Console Mobile App
© 2026, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

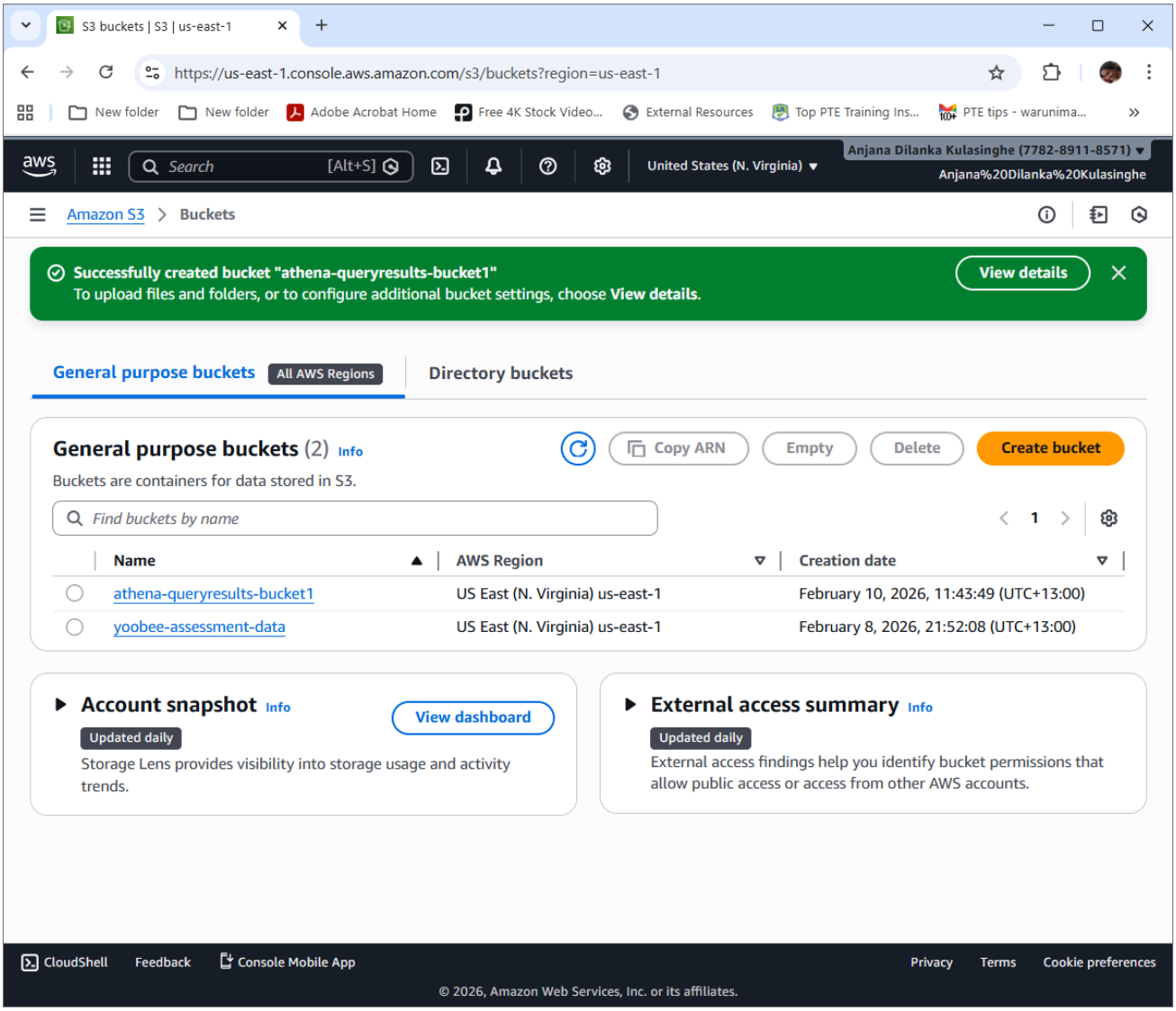

csv_upload_notifier | Function
csv_upload_notifier | Function
CloudWatch | us-east-1
Notify_CSVUploads | Topics
https://us-east-1.console.aws.amazon.com/cloudwatch/home?region=us-east-1#logsV2:log-groups/log-grou...
New Chrome available
New folder
New folder
Adobe Acrobat Home
Free 4K Stock Video...
External Resources
Top PTE Training Ins...
PTE tips - warunima...
Search
[Alt+S]
United States (N. Virginia)
Anjana Dilanka Kulasinghe (7782-8911-8571)
Anjana%20Dilanka%20Kulasinghe
CloudWatch
Log management
/aws/lambda/csv_upload_notifier
CloudWatch
Favorites and recents
Ingestion
Dashboards
Alarms 0 0 0
In alarm
All alarms
Billing
AI Operations
GenAI Observability
Application Signals (APM)
Infrastructure Monitoring
Logs
Log Management New
Log Anomalies
Live Tail
Logs Insights
/aws/lambda/csv_upload_notifier
Actions
View in Logs Insights
Start tailing
Search log group
Log group details
Log class Info
Standard
ARN
arn:aws:logs:us-east-1:778289118571:log-group:/aws/lambda/csv_upload_notifier:*
Creation time
-
Retention
Never expire
Stored bytes
-
Account
-
Metric filters
Subscription filters
Not supported
Contributor Insights rules
Not supported
KMS key ID
-
Deletion protection
Off
Data protection
-
Sensitive data count
Not supported
Custom field indexes
Not supported
Transformer
Not supported
Anomaly detection
Not supported
Log streams
Tags
Data protection
Anomaly detection
Metric filters
Subscri
CloudShell
Feedback
Console Mobile App
© 2026, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

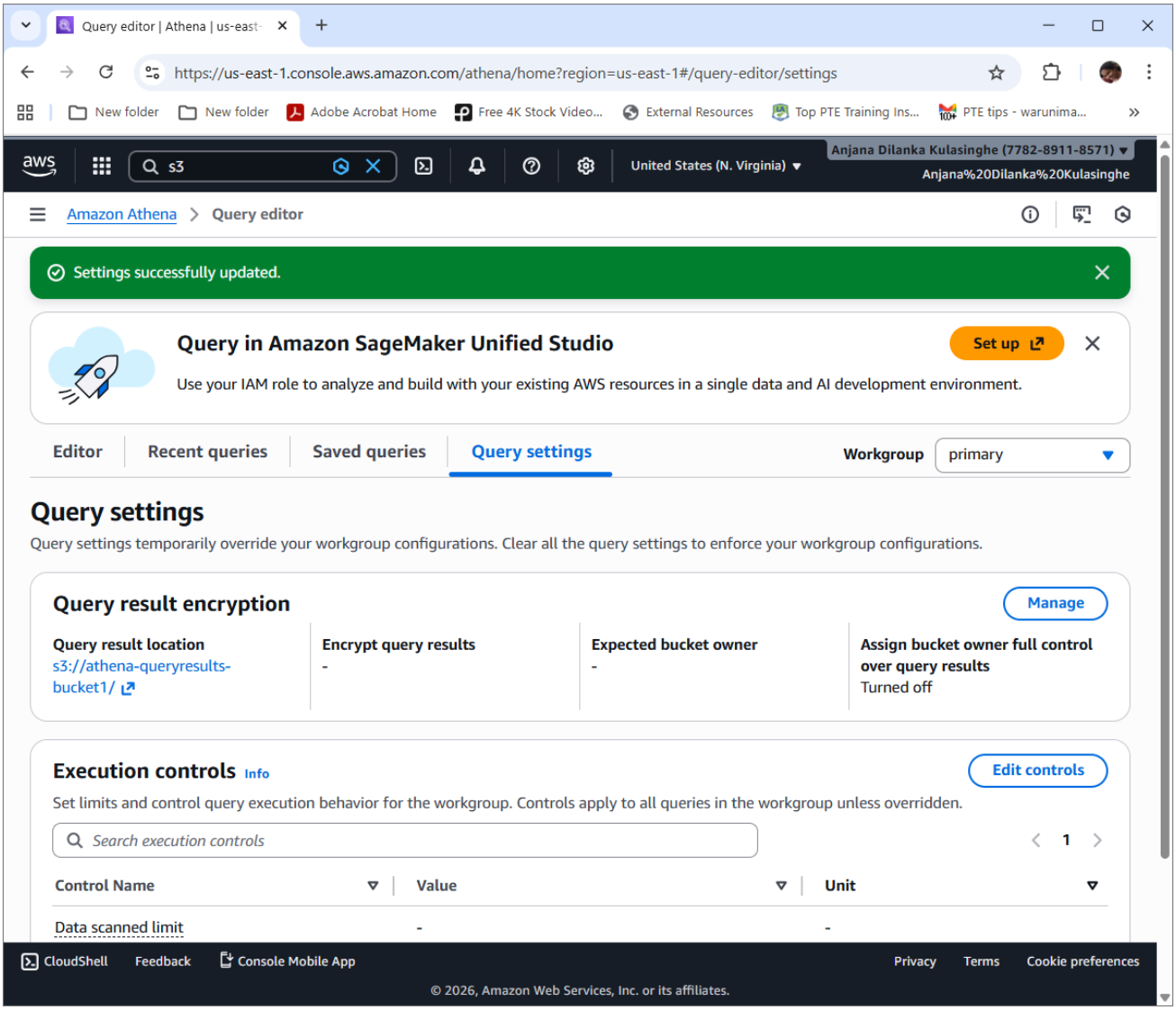
Task 2: Query and Visualize the Dataset Using Athena and QuickSight.

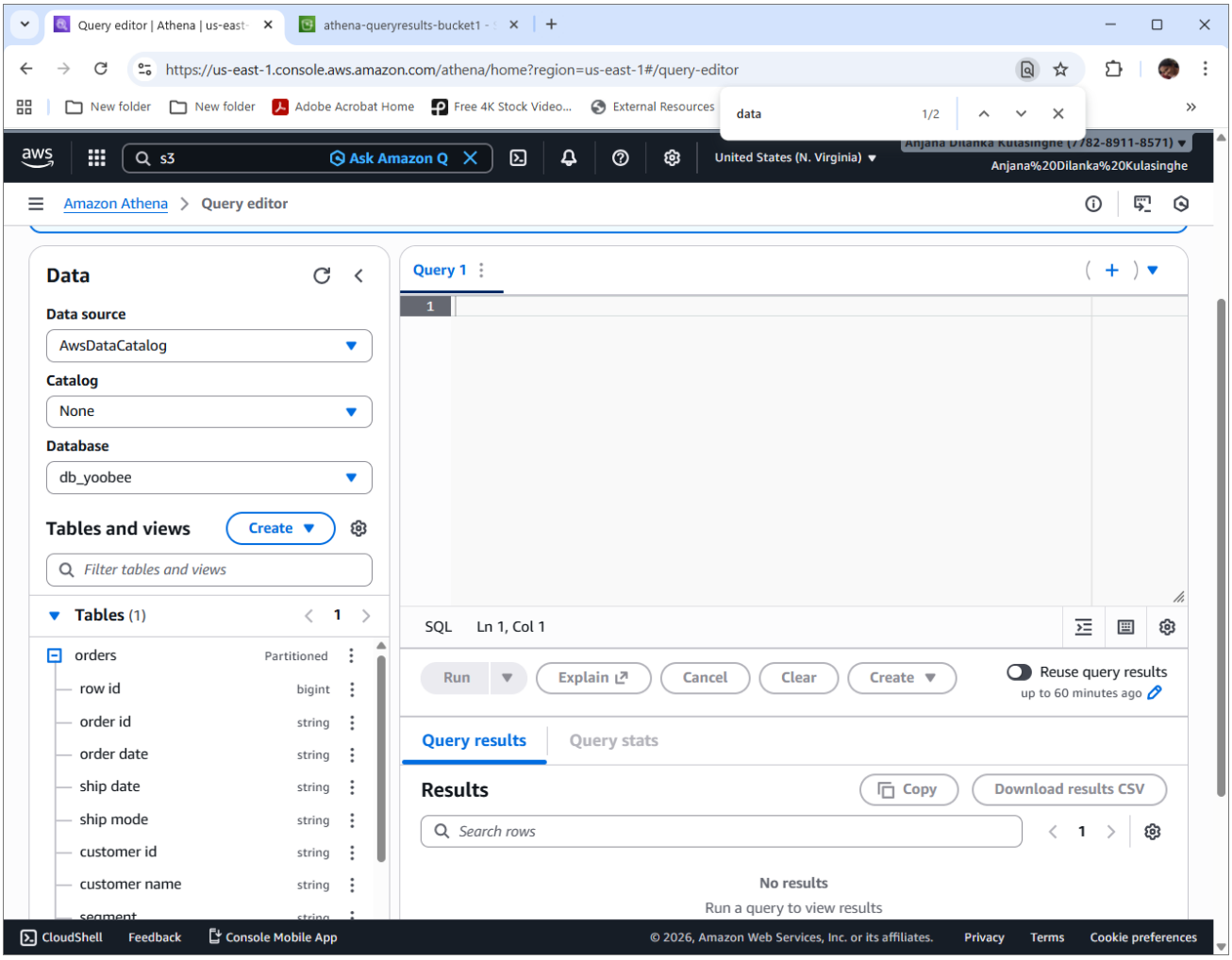
In this task, Amazon Athena was used to query the dataset stored in S3, and Amazon QuickSight was used to visualize analytical insights. This stage transformed raw data into meaningful business intelligence.

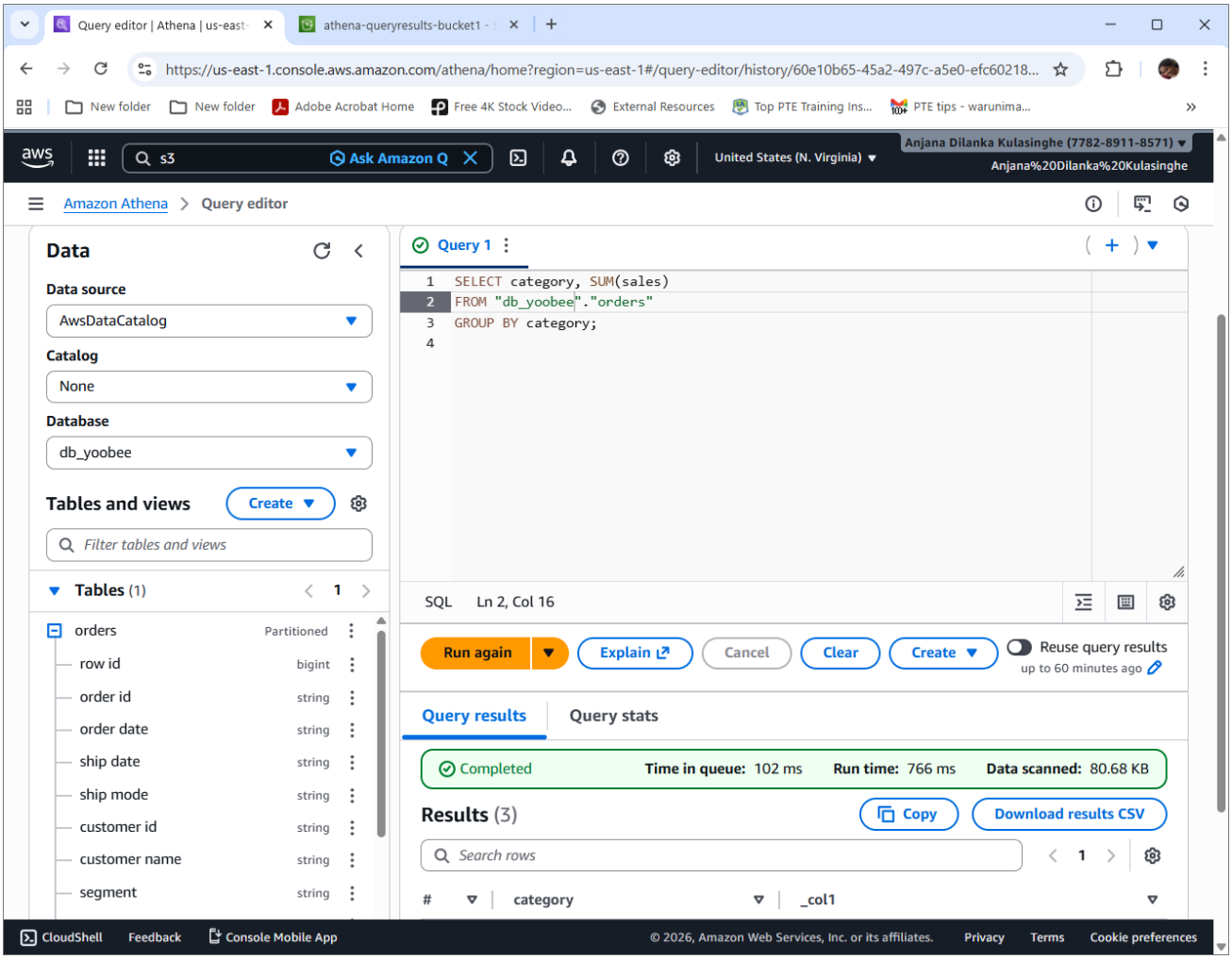
Activities: 2.1: Amazon Athena Queries

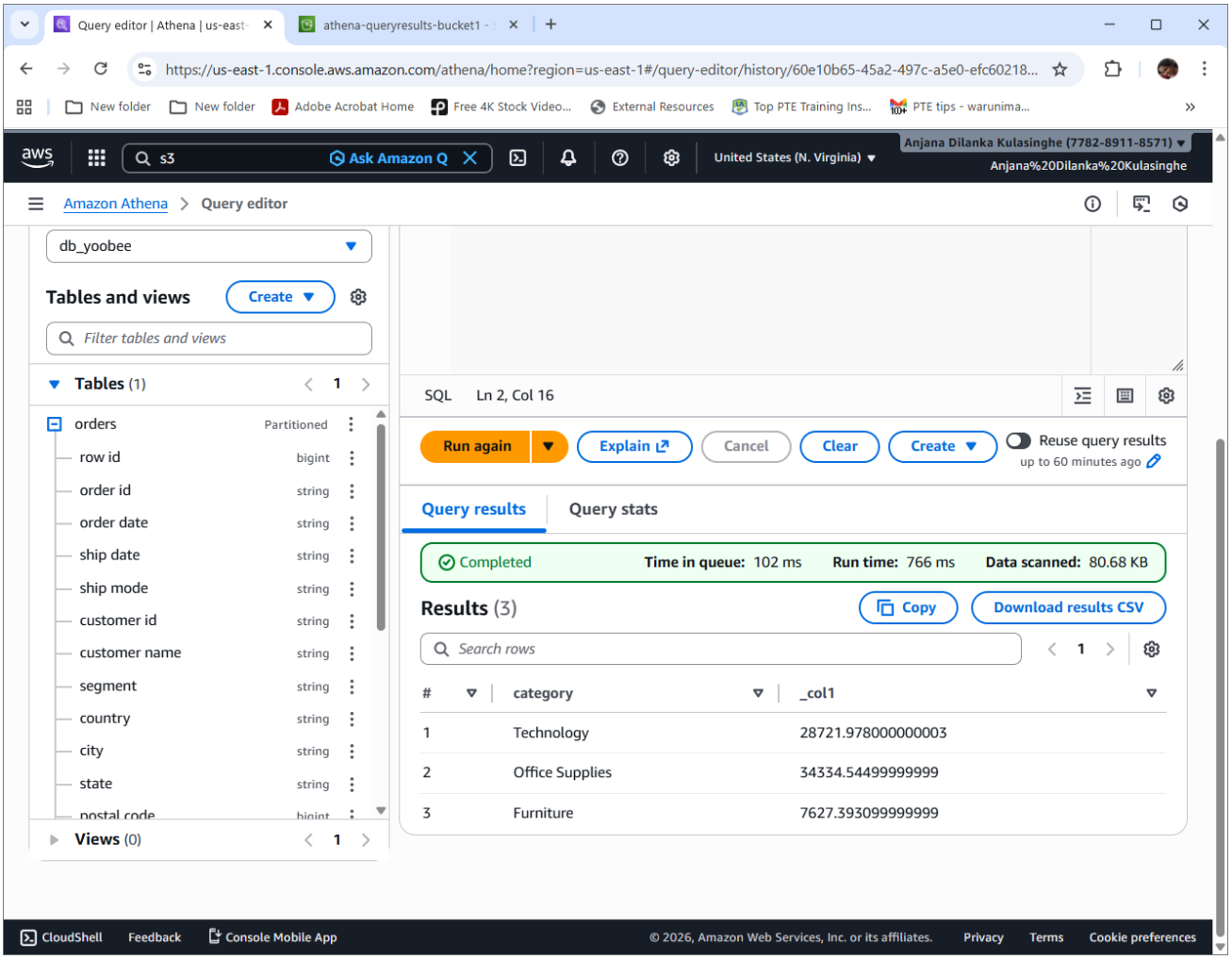
An S3 bucket was configured to store Athena query results. Athena was connected to the Glue Data Catalog, allowing SQL queries to be executed directly on the orders table.
Queries were executed to calculate total sales by category, retrieve full dataset records, and filter by specific snapshot dates. When a second dataset was uploaded and the crawler re-run, the data scanned size increased, which indicated that the additional dataset had been successfully included in the table.

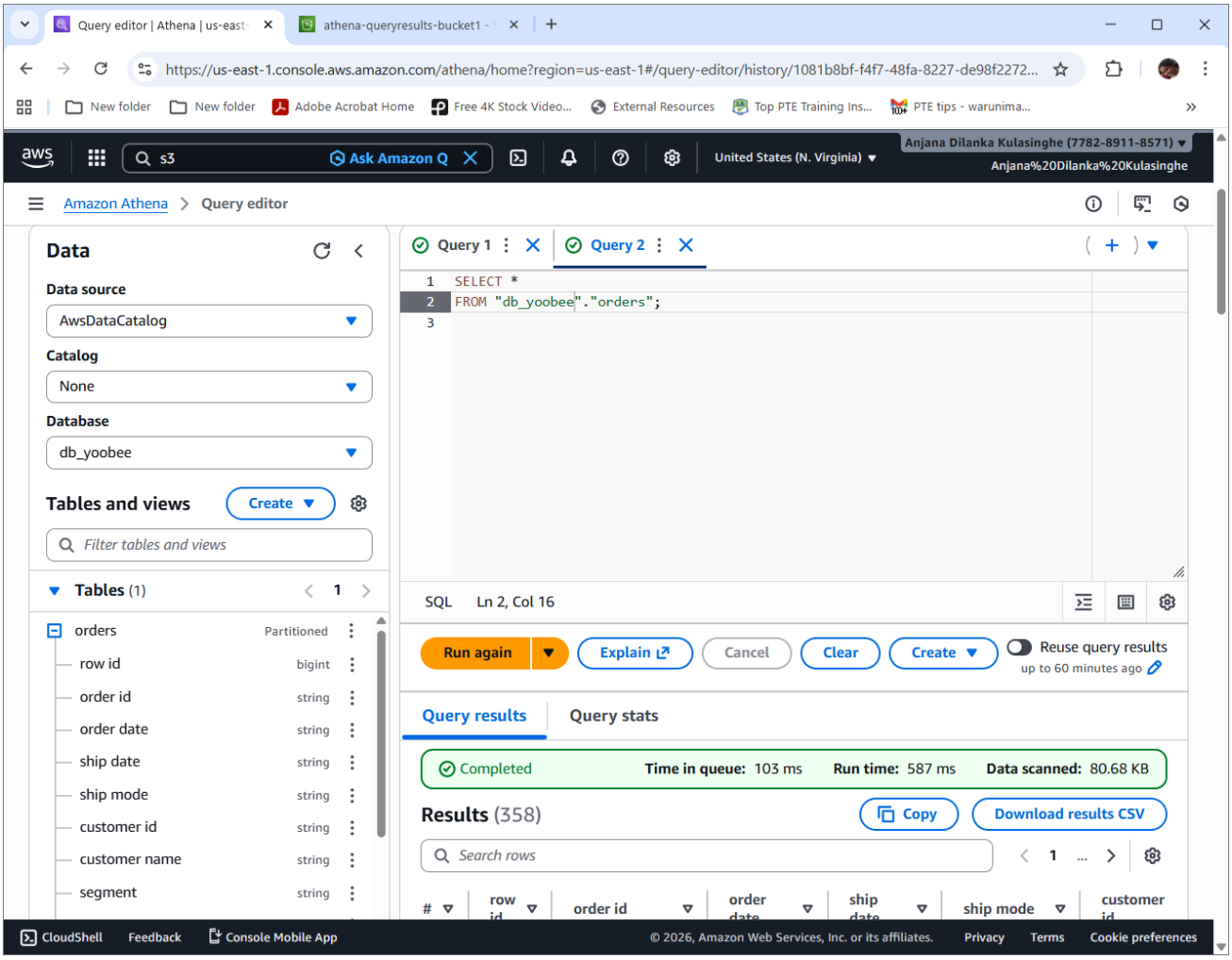


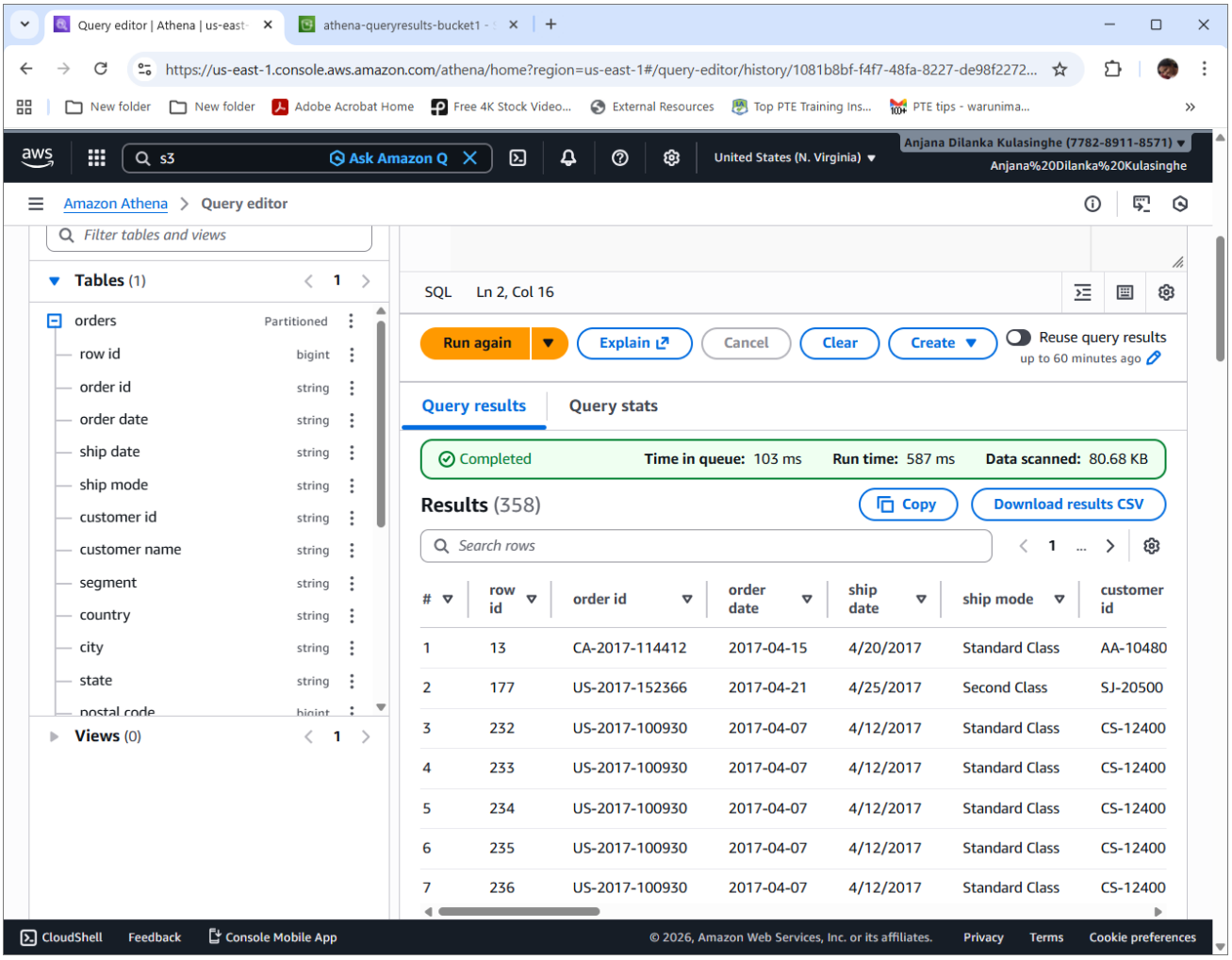

Query editor | Athena | us-east-
https://us-east-1.console.aws.amazon.com/athena/home?region=us-east-1#/query-editor/settings
New folder
New folder
Adobe Acrobat Home
Free 4K Stock Video...
External Resources
Top PTE Training Ins...
PTE tips - warunima...
aws
s3
United States (N. Virginia)
Anjana Dilanka Kulasinghe (7782-8911-8571)
Anjana%20Dilanka%20Kulasinghe
Amazon Athena
Query editor
Settings successfully updated.
Query in Amazon SageMaker Unified Studio
Set up
Use your IAM role to analyze and build with your existing AWS resources in a single data and AI development environment.
Editor
Recent queries
Saved queries
Query settings
Workgroup
primary
Query settings
Query settings temporarily override your workgroup configurations. Clear all the query settings to enforce your workgroup configurations.
Query result encryption
Manage
Query result location
s3://athena-queryresults-bucket1/
Encrypt query results
-
Expected bucket owner
-
Assign bucket owner full control over query results
Turned off
Execution controls
Info
Edit controls
Set limits and control query execution behavior for the workgroup. Controls apply to all queries in the workgroup unless overridden.
Search execution controls
1
Control Name
Value
Unit
Data scanned limit
-
-
CloudShell
Feedback
Console Mobile App
Privacy
Terms
Cookie preferences
© 2026, Amazon Web Services, Inc. or its affiliates.

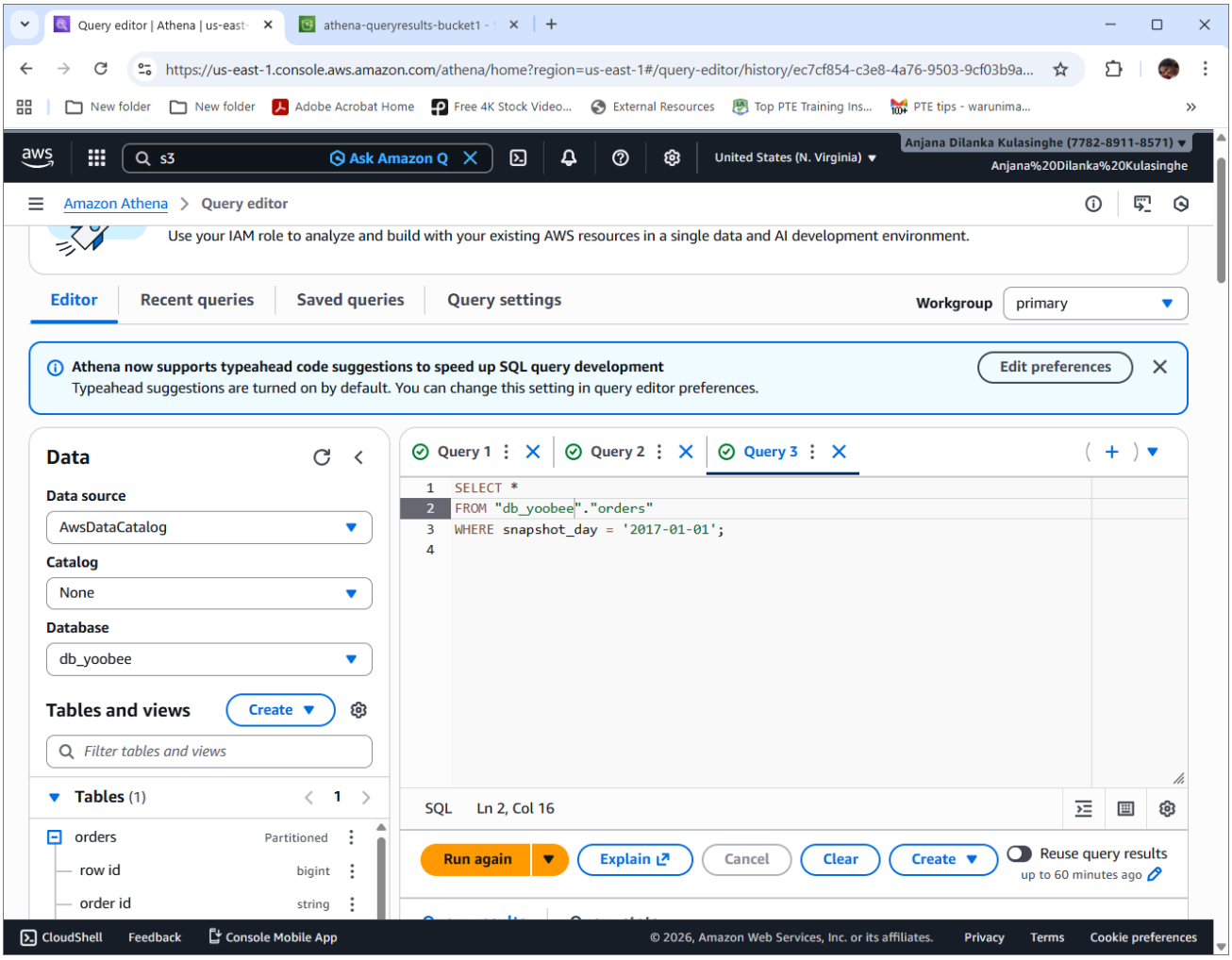

Query editor | Athena | us-east-
athena-queryresults-bucket1 -
https://us-east-1.console.aws.amazon.com/athena/home?region=us-east-1#/query-editor
New folder
New folder
Adobe Acrobat Home
Free 4K Stock Video...
External Resources
data
1/2
aws
s3
Ask Amazon Q
United States (N. Virginia)
Anjana%20Dilanka%20Kulasinghe
Amazon Athena
Query editor
Data
Data source
AwsDataCatalog
Catalog
None
Database
db_yoobee
Tables and views
Create
Filter tables and views
Tables (1)
orders
Partitioned
row id
bigint
order id
string
order date
string
ship date
string
ship mode
string
customer id
string
customer name
string
Query 1
SQL
Ln 1, Col 1
Run
Explain
Cancel
Clear
Create
Reuse query results
up to 60 minutes ago
Query results
Query stats
Results
Copy
Download results CSV
Search rows
No results
Run a query to view results
CloudShell
Feedback
Console Mobile App
© 2026, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

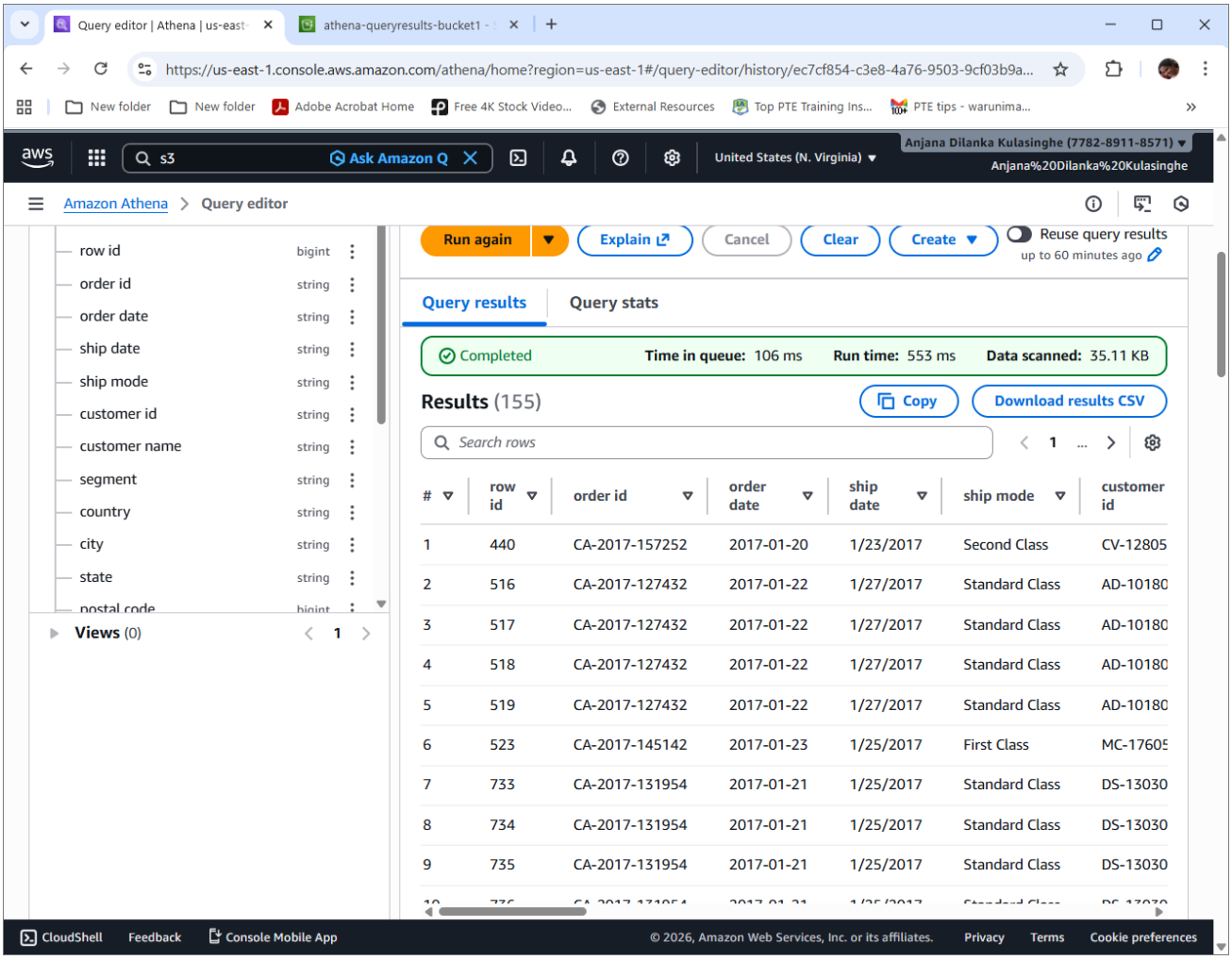

Query editor | Athena | us-east-
athena-queryresults-bucket1 -
https://us-east-1.console.aws.amazon.com/athena/home?region=us-east-1#/query-editor/history/60e10b65-45a2-497c-a5e0-efc60218...
New folder
New folder
Adobe Acrobat Home
Free 4K Stock Video...
External Resources
Top PTE Training Ins...
PTE tips - warunima...
aws
s3
Ask Amazon Q
United States (N. Virginia)
Anjana Dilanka Kulasinghe (7782-8911-8571)
Anjana%20Dilanka%20Kulasinghe
Amazon Athena
Query editor
Data
Data source
AwsDataCatalog
Catalog
None
Database
db_yoobee
Tables and views
Create
Filter tables and views
Tables (1)
orders
Partitioned
row id
bigint
order id
string
order date
string
ship date
string
ship mode
string
customer id
string
customer name
string
segment
string
Query 1
1 SELECT category, SUM(sales)
2 FROM "db_yoobee"."orders"
3 GROUP BY category;
4
SQL
Ln 2, Col 16
Run again
Explain
Cancel
Clear
Create
Reuse query results
up to 60 minutes ago
Query results
Query stats
Completed
Time in queue: 102 ms
Run time: 766 ms
Data scanned: 80.68 KB
Results (3)
Copy
Download results CSV
Search rows
#
category
_col1
CloudShell
Feedback
Console Mobile App
© 2026, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences


Query editor | Athena | us-east-
athena-queryresults-bucket1 -
https://us-east-1.console.aws.amazon.com/athena/home?region=us-east-1#/query-editor/history/60e10b65-45a2-497c-a5e0-efc60218...
New folder
New folder
Adobe Acrobat Home
Free 4K Stock Video...
External Resources
Top PTE Training Ins...
PTE tips - warunima...
aws
s3
Ask Amazon Q
United States (N. Virginia)
Anjana Dilanka Kulasinghe (7782-8911-8571)
Anjana%20Dilanka%20Kulasinghe
Amazon Athena
Query editor
db_yoobee
Tables and views
Create
Filter tables and views
Tables (1)
orders
Partitioned
row id
bigint
order id
string
order date
string
ship date
string
ship mode
string
customer id
string
customer name
string
segment
string
country
string
city
string
state
string
Views (0)
SQL
Ln 2, Col 16
Run again
Explain
Cancel
Clear
Create
Reuse query results
up to 60 minutes ago
Query results
Query stats
Completed
Time in queue: 102 ms
Run time: 766 ms
Data scanned: 80.68 KB
Results (3)
Copy
Download results CSV
Search rows
category _col1
1 Technology 28721.978000000003
2 Office Supplies 34334.54499999999
3 Furniture 7627.393099999999
CloudShell
Feedback
Console Mobile App
© 2026, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences


Query editor | Athena | us-east-
athena-queryresults-bucket1 -
https://us-east-1.console.aws.amazon.com/athena/home?region=us-east-1#/query-editor/history/1081b8bf-f4f7-48fa-8227-de98f2272...
New folder
New folder
Adobe Acrobat Home
Free 4K Stock Video...
External Resources
Top PTE Training Ins...
PTE tips - warunima...
aws
s3
Ask Amazon Q
United States (N. Virginia)
Anjana Dilanka Kulasinghe (7782-8911-8571)
Anjana%20Dilanka%20Kulasinghe
Amazon Athena
Query editor
Data
Data source
AwsDataCatalog
Catalog
None
Database
db_yoobee
Tables and views
Create
Filter tables and views
Tables (1)
orders
Partitioned
row id
bigint
order id
string
order date
string
ship date
string
ship mode
string
customer id
string
customer name
string
segment
string
Query 1
Query 2
1 SELECT *
2 FROM "db_yoobee"."orders";
3
SQL
Ln 2, Col 16
Run again
Explain
Cancel
Clear
Create
Reuse query results
up to 60 minutes ago
Query results
Query stats
Completed
Time in queue: 103 ms
Run time: 587 ms
Data scanned: 80.68 KB
Results (358)
Copy
Download results CSV
Search rows
1
...
#
row id
order id
order date
ship date
ship mode
customer id
CloudShell
Feedback
Console Mobile App
© 2026, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences


Query editor | Athena | us-east-
athena-queryresults-bucket1 -
https://us-east-1.console.aws.amazon.com/athena/home?region=us-east-1#/query-editor/history/1081b8bf-f4f7-48fa-8227-de98f2272...
New folder
New folder
Adobe Acrobat Home
Free 4K Stock Video...
External Resources
Top PTE Training Ins...
PTE tips - warunima...
aws
s3
Ask Amazon Q
United States (N. Virginia)
Anjana Dilanka Kulasinghe (7782-8911-8571)
Anjana%20Dilanka%20Kulasinghe
Amazon Athena
Query editor
Filter tables and views
Tables (1)
orders
Partitioned
row id
bigint
order id
string
order date
string
ship date
string
ship mode
string
customer id
string
customer name
string
segment
string
country
string
city
string
state
string
Views (0)
SQL
Ln 2, Col 16
Run again
Explain
Cancel
Clear
Create
Reuse query results
up to 60 minutes ago
Query results
Query stats
Completed
Time in queue: 103 ms
Run time: 587 ms
Data scanned: 80.68 KB
Results (358)
Copy
Download results CSV
Search rows
| row id | order id | order date | ship date | ship mode | customer id
1 | 13 | CA-2017-114412 | 2017-04-15 | 4/20/2017 | Standard Class | AA-10480
2 | 177 | US-2017-152366 | 2017-04-21 | 4/25/2017 | Second Class | SJ-20500
3 | 232 | US-2017-100930 | 2017-04-07 | 4/12/2017 | Standard Class | CS-12400
4 | 233 | US-2017-100930 | 2017-04-07 | 4/12/2017 | Standard Class | CS-12400
5 | 234 | US-2017-100930 | 2017-04-07 | 4/12/2017 | Standard Class | CS-12400
6 | 235 | US-2017-100930 | 2017-04-07 | 4/12/2017 | Standard Class | CS-12400
7 | 236 | US-2017-100930 | 2017-04-07 | 4/12/2017 | Standard Class | CS-12400
CloudShell
Feedback
Console Mobile App
© 2026, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences


Query editor | Athena | us-east-
athena-queryresults-bucket1 -
https://us-east-1.console.aws.amazon.com/athena/home?region=us-east-1#/query-editor/history/ec7cf854-c3e8-4a76-9503-9cf03b9a...
New folder
New folder
Adobe Acrobat Home
Free 4K Stock Video...
External Resources
Top PTE Training Ins...
PTE tips - warunima...
aws
s3
Ask Amazon Q
United States (N. Virginia)
Anjana Dilanka Kulasinghe (7782-8911-8571)
Anjana%20Dilanka%20Kulasinghe
Amazon Athena
Query editor
Use your IAM role to analyze and build with your existing AWS resources in a single data and AI development environment.
Editor
Recent queries
Saved queries
Query settings
Workgroup
primary
Athena now supports typeahead code suggestions to speed up SQL query development
Typeahead suggestions are turned on by default. You can change this setting in query editor preferences.
Edit preferences
Data
Data source
AwsDataCatalog
Catalog
None
Database
db_yoobee
Tables and views
Create
Filter tables and views
Tables (1)
orders
Partitioned
row id
bigint
order id
string
Query 1
Query 2
Query 3
1 SELECT *
2 FROM "db_yoobee"."orders"
3 WHERE snapshot_day = '2017-01-01';
4
SQL
Ln 2, Col 16
Run again
Explain
Cancel
Clear
Create
Reuse query results
up to 60 minutes ago
CloudShell
Feedback
Console Mobile App
© 2026, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

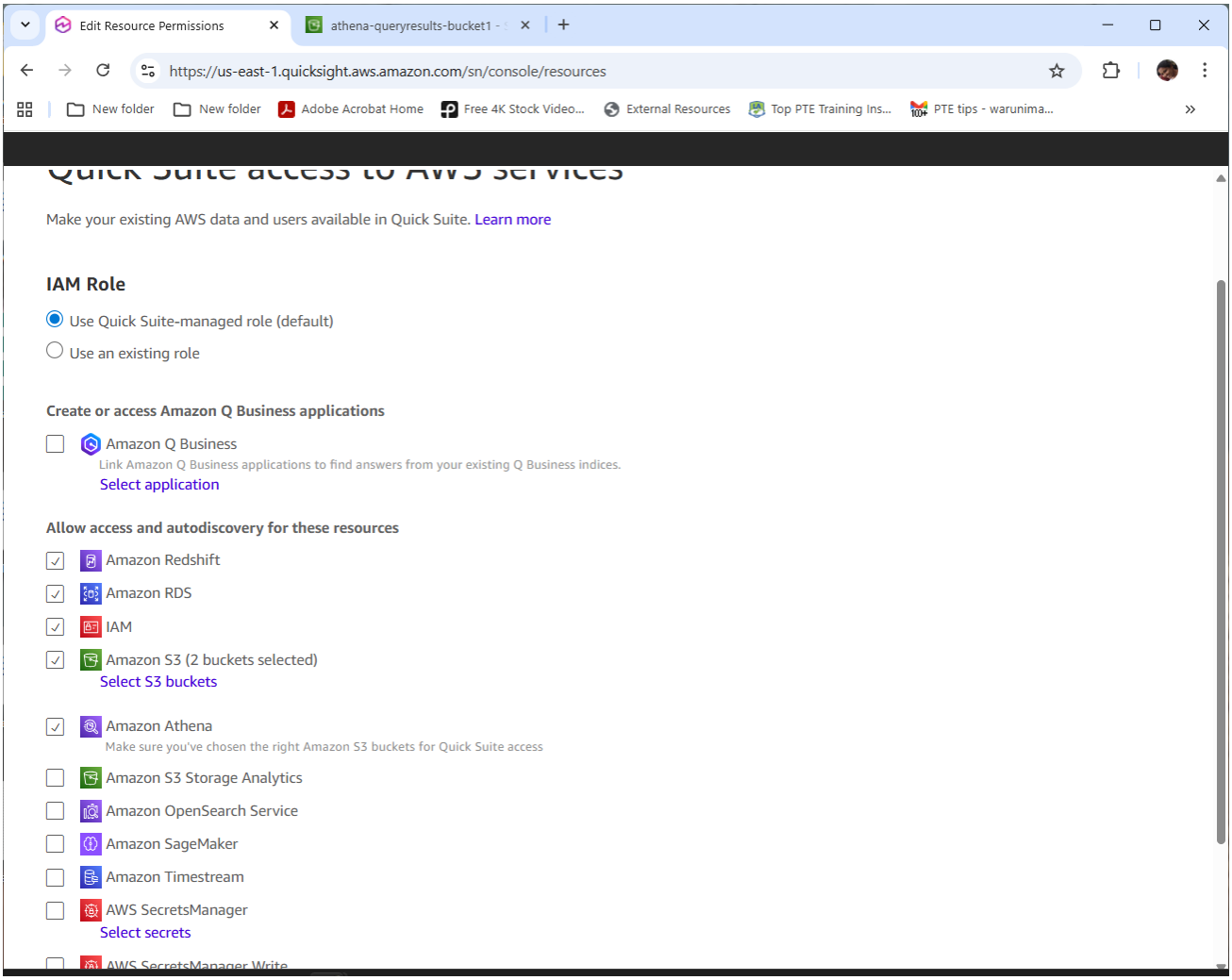


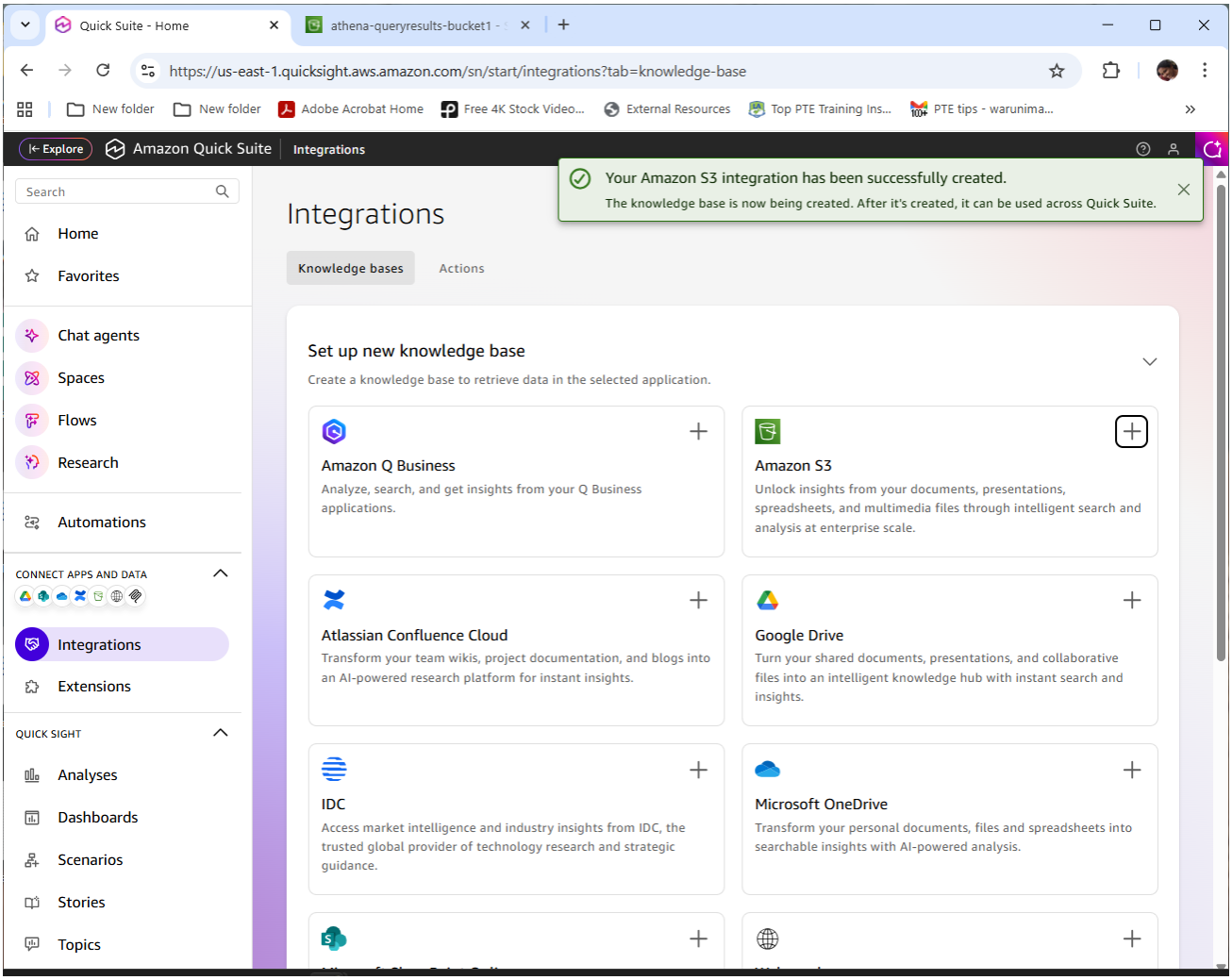
Activities: 2.2: Amazon QuickSight Dashboards

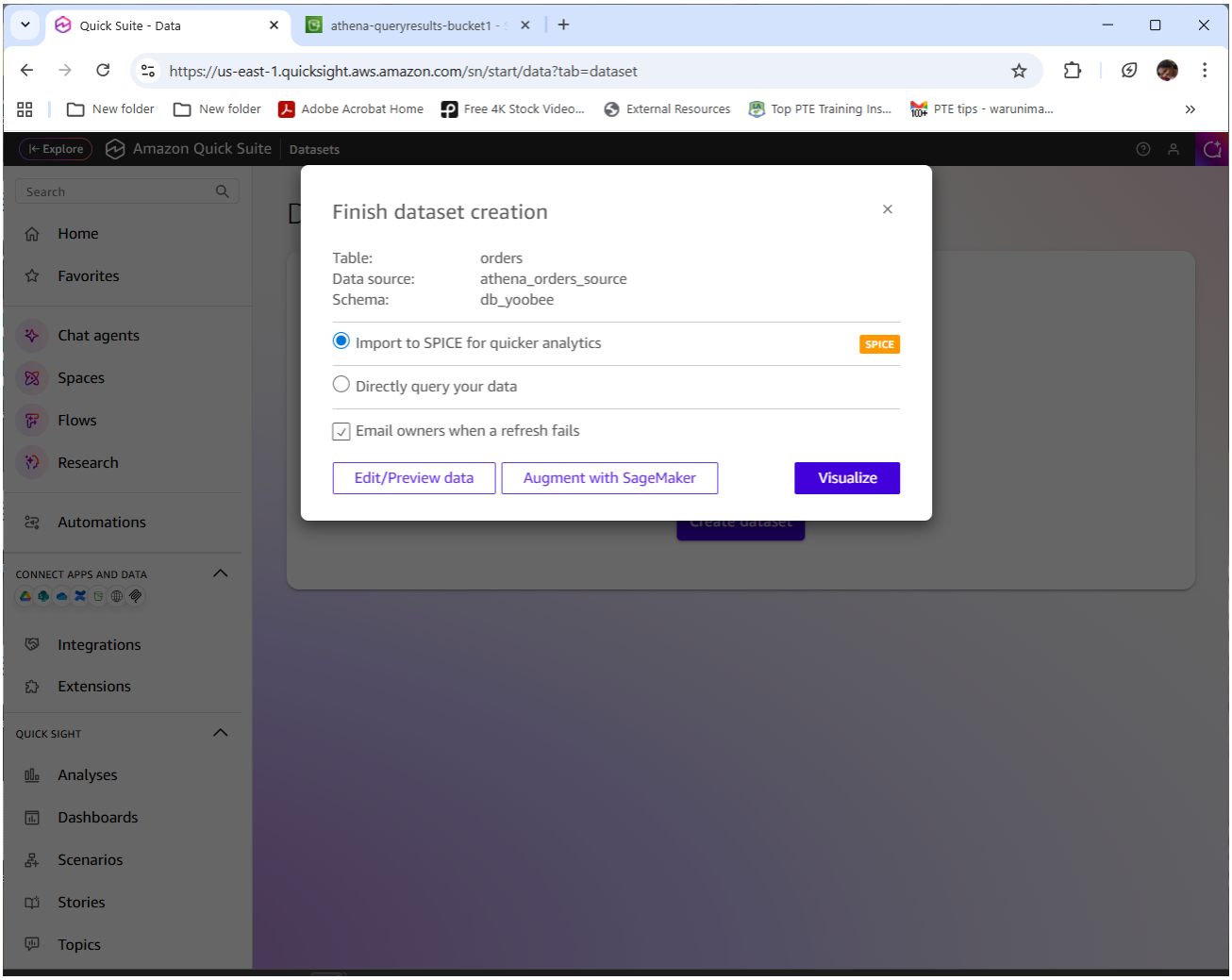
Amazon QuickSight was configured using the Standard Edition and connected to Athena. SPICE was used to import the dataset for faster visual performance.

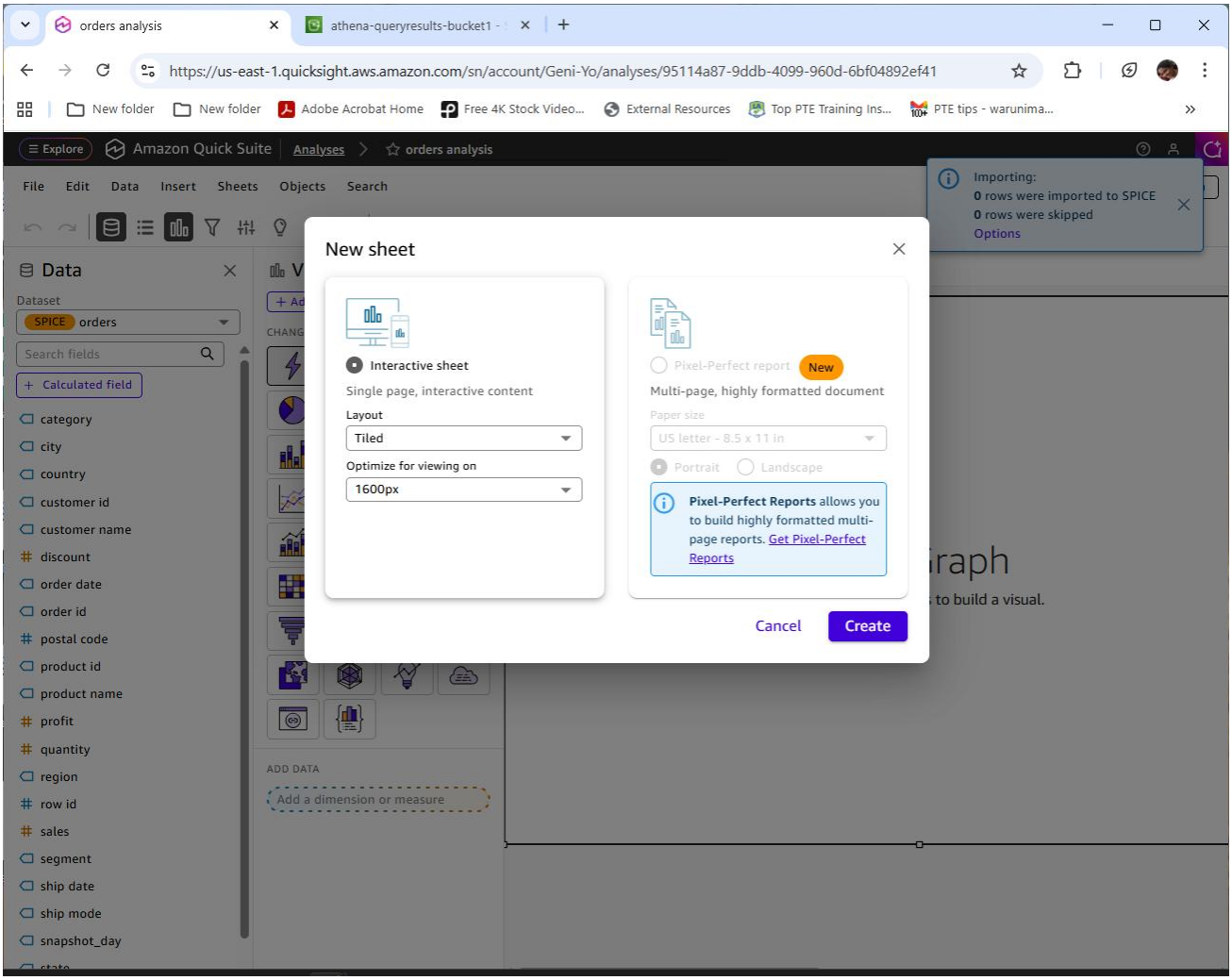
From the chart, it was clear that certain categories generated significantly higher sales, which would help decision-makers prioritise products.

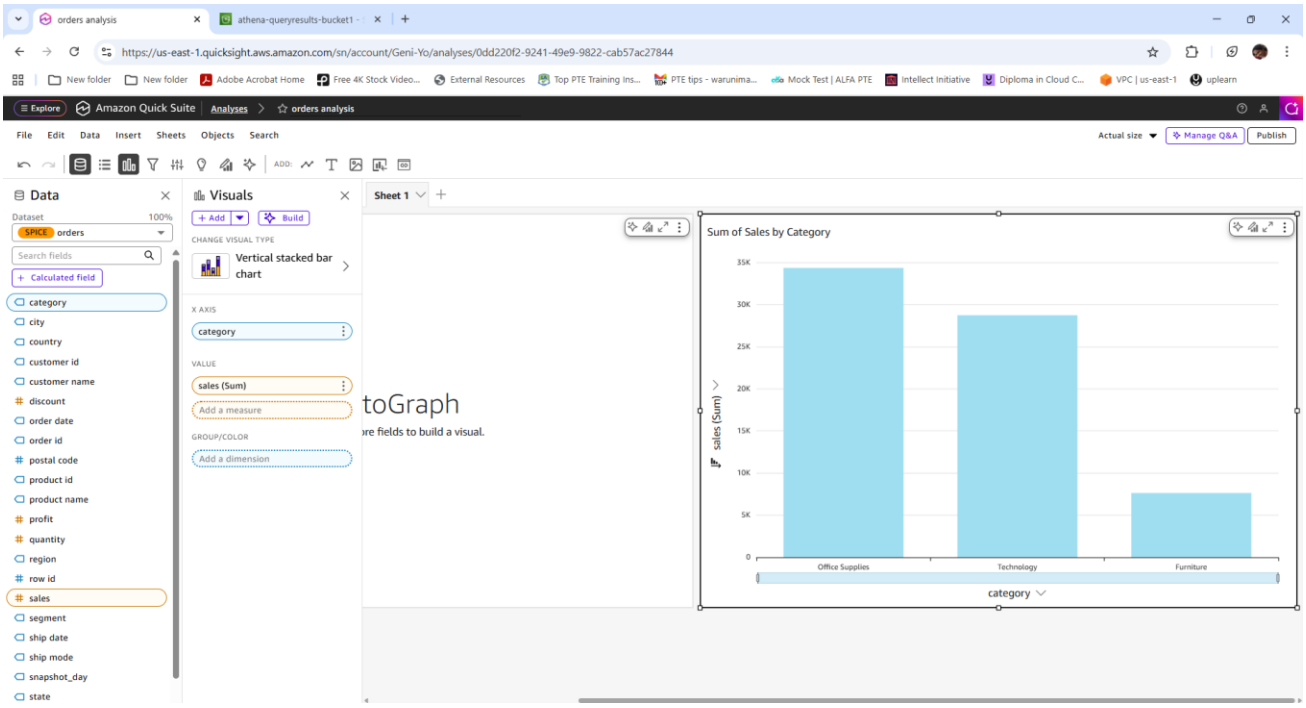
A bar chart was created to show Sales by Category, and another visual was built to represent Profit by City. These dashboards transformed numerical results into visual insights that would support business decision-making.

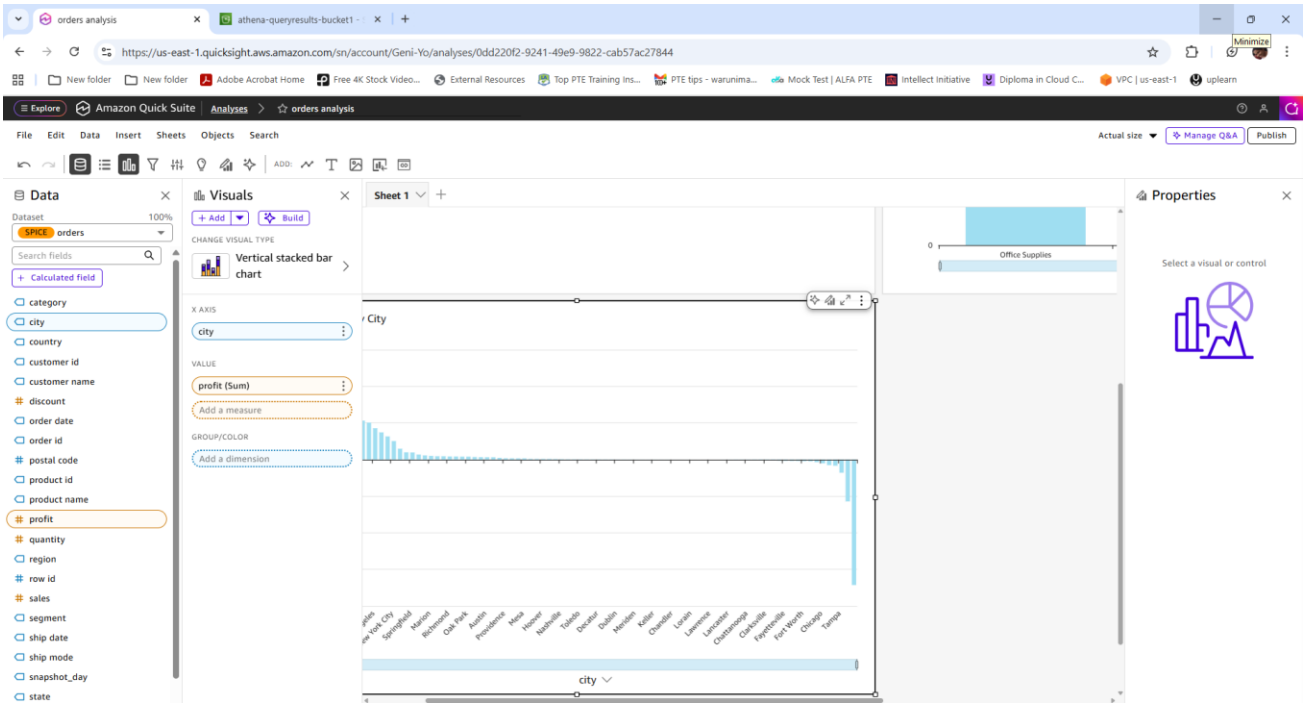

Edit Resource Permissions
athena-queryresults-bucket1 -
https://us-east-1.quicksight.aws.amazon.com/sn/console/resources
New folder
New folder
Adobe Acrobat Home
Free 4K Stock Video...
External Resources
Top PTE Training Ins...
PTE tips - warunima...
Make your existing AWS data and users available in Quick Suite. Learn more
IAM Role
Use Quick Suite-managed role (default)
Use an existing role
Create or access Amazon Q Business applications
Amazon Q Business
Link Amazon Q Business applications to find answers from your existing Q Business indices.
Select application
Allow access and autodiscovery for these resources
Amazon Redshift
Amazon RDS
IAM
Amazon S3 (2 buckets selected)
Select S3 buckets
Amazon Athena
Make sure you've chosen the right Amazon S3 buckets for Quick Suite access
Amazon S3 Storage Analytics
Amazon OpenSearch Service
Amazon SageMaker
Amazon Timestream
AWS SecretsManager
Select secrets


Quick Suite - Home
athena-queryresults-bucket1 -
https://us-east-1.quicksight.aws.amazon.com/sn/start/integrations?tab=knowledge-base
New folder
New folder
Adobe Acrobat Home
Free 4K Stock Video...
External Resources
Top PTE Training Ins...
PTE tips - warunima...
Explore
Amazon Quick Suite
Integrations
Search
Home
Favorites
Chat agents
Spaces
Flows
Research
Automations
CONNECT APPS AND DATA
Integrations
Extensions
QUICK SIGHT
Analyses
Dashboards
Scenarios
Stories
Topics
Your Amazon S3 integration has been successfully created.
The knowledge base is now being created. After it's created, it can be used across Quick Suite.
Integrations
Knowledge bases
Actions
Set up new knowledge base
Create a knowledge base to retrieve data in the selected application.
Amazon Q Business
Analyze, search, and get insights from your Q Business applications.
Amazon S3
Unlock insights from your documents, presentations, spreadsheets, and multimedia files through intelligent search and analysis at enterprise scale.
Atlassian Confluence Cloud
Transform your team wikis, project documentation, and blogs into an AI-powered research platform for instant insights.
Google Drive
Turn your shared documents, presentations, and collaborative files into an intelligent knowledge hub with instant search and insights.
IDC
Access market intelligence and industry insights from IDC, the trusted global provider of technology research and strategic guidance.
Microsoft OneDrive
Transform your personal documents, files and spreadsheets into searchable insights with AI-powered analysis.


Quick Suite - Data
athena-queryresults-bucket1 -
https://us-east-1.quicksight.aws.amazon.com/sn/start/data?tab=dataset
New folder
New folder
Adobe Acrobat Home
Free 4K Stock Video...
External Resources
Top PTE Training Ins...
PTE tips - warunima...
Explore
Amazon Quick Suite
Datasets
Search
Home
Favorites
Chat agents
Spaces
Flows
Research
Automations
CONNECT APPS AND DATA
Integrations
Extensions
QUICK SIGHT
Analyses
Dashboards
Scenarios
Stories
Topics
Finish dataset creation
Table: orders
Data source: athena_orders_source
Schema: db_yoobee
Import to SPICE for quicker analytics
SPICE
Directly query your data
Email owners when a refresh fails
Edit/Preview data
Augment with SageMaker
Visualize


orders analysis
athena-queryresults-bucket1 -
https://us-east-1.quicksight.aws.amazon.com/sn/account/Geni-Yo/analyses/95114a87-9ddb-4099-960d-6bf04892ef41
Amazon Quick Suite
Analyses
orders analysis
File Edit Data Insert Sheets Objects Search
Importing:
0 rows were imported to SPICE
0 rows were skipped
Options
Data
Dataset
SPICE orders
Search fields
Calculated field
category
city
country
customer id
customer name
discount
order date
order id
postal code
product id
product name
profit
quantity
region
row id
sales
segment
ship date
ship mode
snapshot_day
ADD DATA
Add a dimension or measure
New sheet
Interactive sheet
Single page, interactive content
Layout
Tiled
Optimize for viewing on
1600px
Pixel-Perfect report
New
Multi-page, highly formatted document
Paper size
US letter - 8.5 x 11 in
Portrait
Landscape
Pixel-Perfect Reports allows you to build highly formatted multi-page reports. Get Pixel-Perfect Reports
Cancel
Create


Visuals
Sheet 1
Vertical stacked bar chart
X AXIS
category
VALUE
sales (Sum)
Add a measure
GROUP/COLOR
Add a dimension
Sum of Sales by Category
sales (Sum)
Office Supplies
Technology
Furniture
category

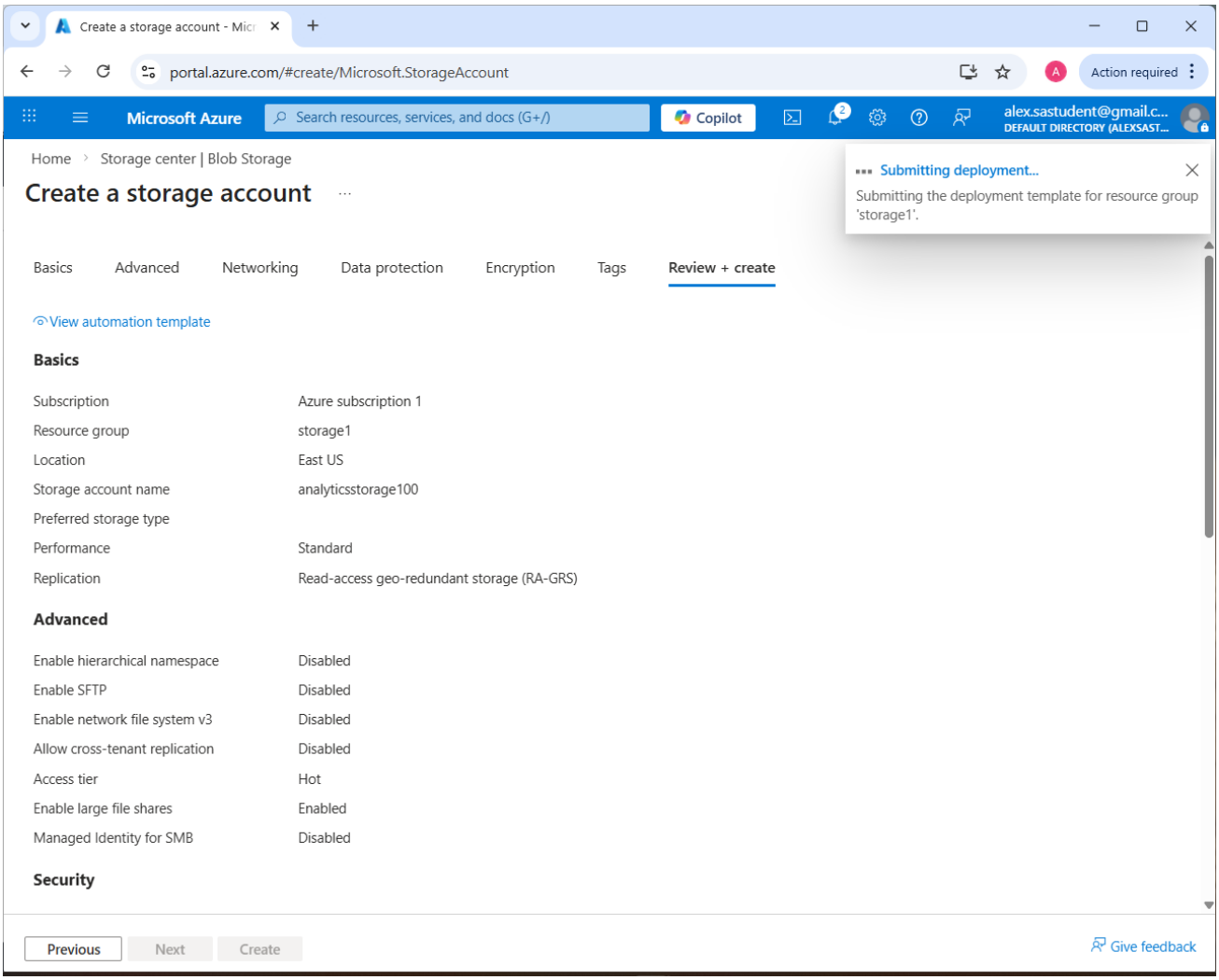

orders analysis
athena-queryresults-bucket1 -
https://us-east-1.quicksight.aws.amazon.com/sn/account/Geni-Yo/analyses/0dd220f2-9241-49e9-9822-cab57ac27844
Amazon Quick Suite
Analyses
orders analysis
File Edit Data Insert Sheets Objects Search
Actual size
Manage Q&A
Publish
Data
Dataset
100%
SPICE orders
Search fields
Calculated field
category
city
country
customer id
customer name
discount
order date
order id
postal code
product id
product name
profit
quantity
region
row id
sales
segment
ship date
ship mode
snapshot_day
state
Visuals
Add
Build
CHANGE VISUAL TYPE
Vertical stacked bar chart
X AXIS
city
VALUE
profit (Sum)
Add a measure
GROUP/COLOR
Add a dimension
Sheet 1
City
Office Supplies
city
Properties
Select a visual or control

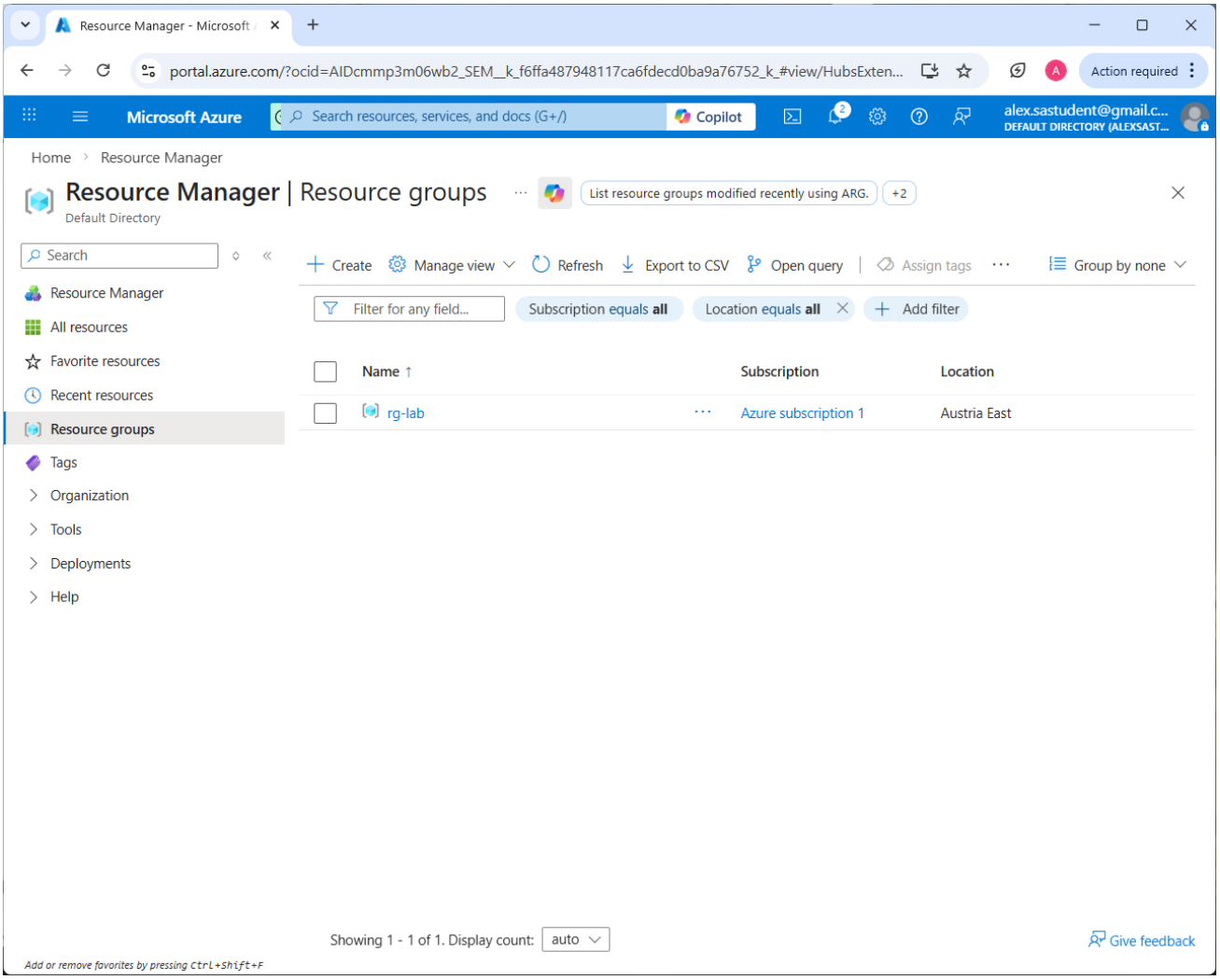
Task 3: Multi-Cloud Integration (GitHub → AWS → Azure Synapse Analytics)

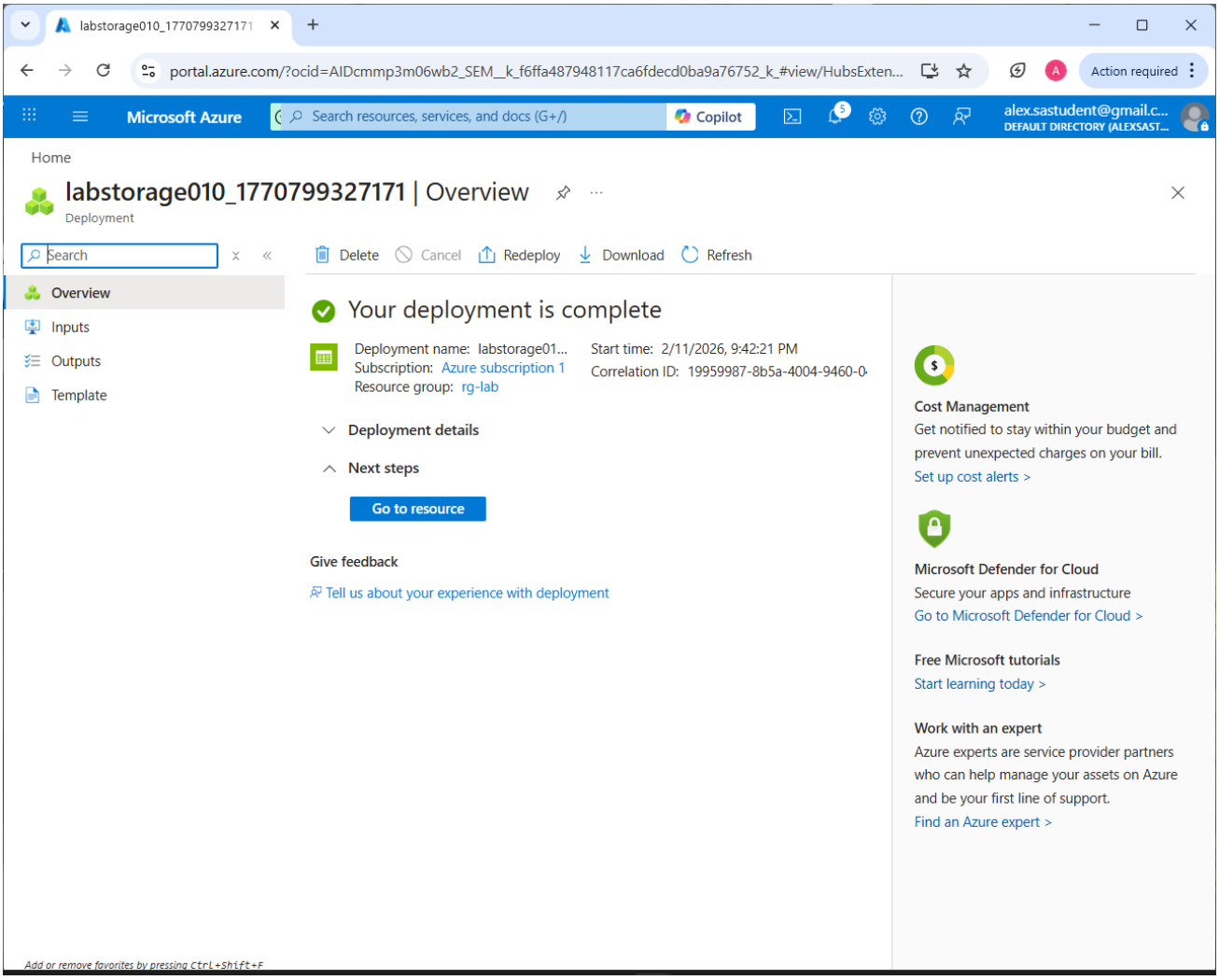
This task demonstrates cross-cloud interoperability by exporting curated AWS query results and analysing them within Microsoft Azure Synapse Analytics. This reflects real-world enterprise scenarios where organisations operate across multiple cloud providers.

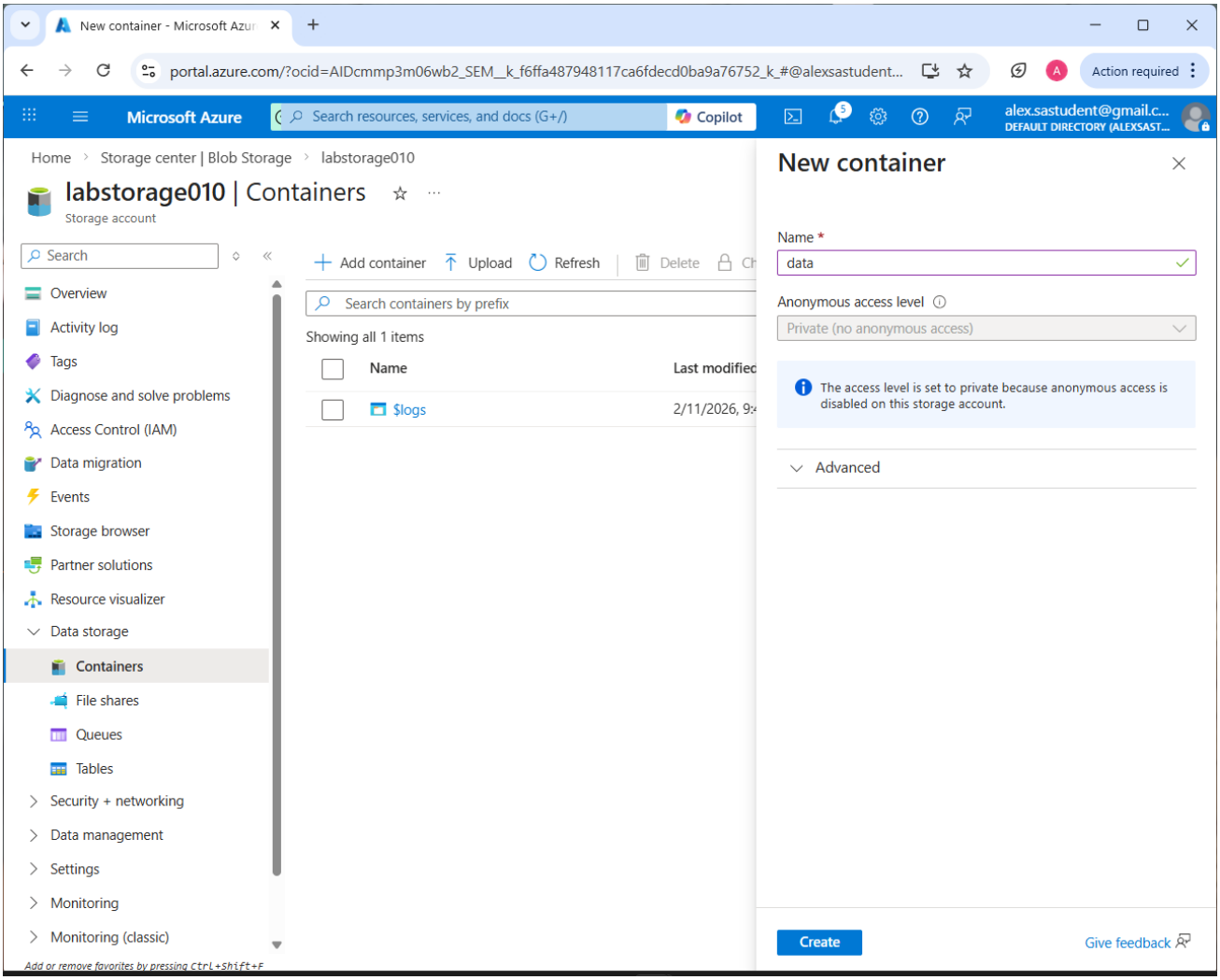
Activity 3.1: Data Export from AWS S3

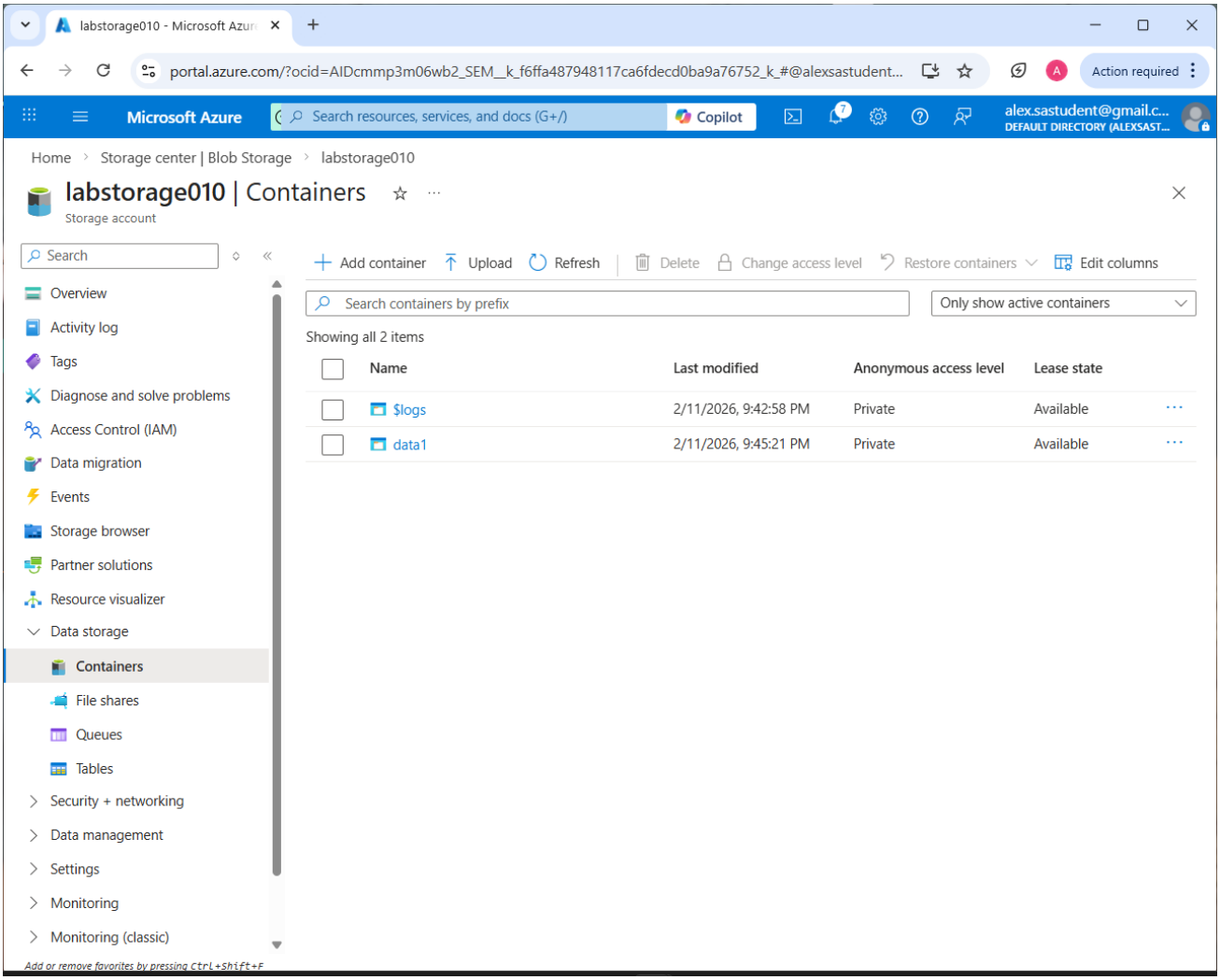
The Athena query results were downloaded from the S3 result location and renamed for clarity. These CSV files represented curated datasets ready for cross-cloud integration.




Resource Manager - Microsoft
portal.azure.com/?ocid=AIDcmmp3m06wb2_SEM__k_f6ffa487948117ca6fdecd0ba9a76752_k_#view/HubsExten...
Action required
Microsoft Azure
Search resources, services, and docs (G+/)
Copilot
alex.sastudent@gmail.c...
DEFAULT DIRECTORY (ALEXSAST...
Home
Resource Manager
Resource Manager | Resource groups
Default Directory
List resource groups modified recently using ARG.
+2
Search
Resource Manager
All resources
Favorite resources
Recent resources
Resource groups
Tags
Organization
Tools
Deployments
Help
Create
Manage view
Refresh
Export to CSV
Open query
Assign tags
Group by none
Filter for any field...
Subscription equals all
Location equals all
Add filter
Name
Subscription
Location
rg-lab
Azure subscription 1
Austria East
Showing 1 - 1 of 1. Display count:
auto
Give feedback
Add or remove favorites by pressing Ctrl+Shift+F

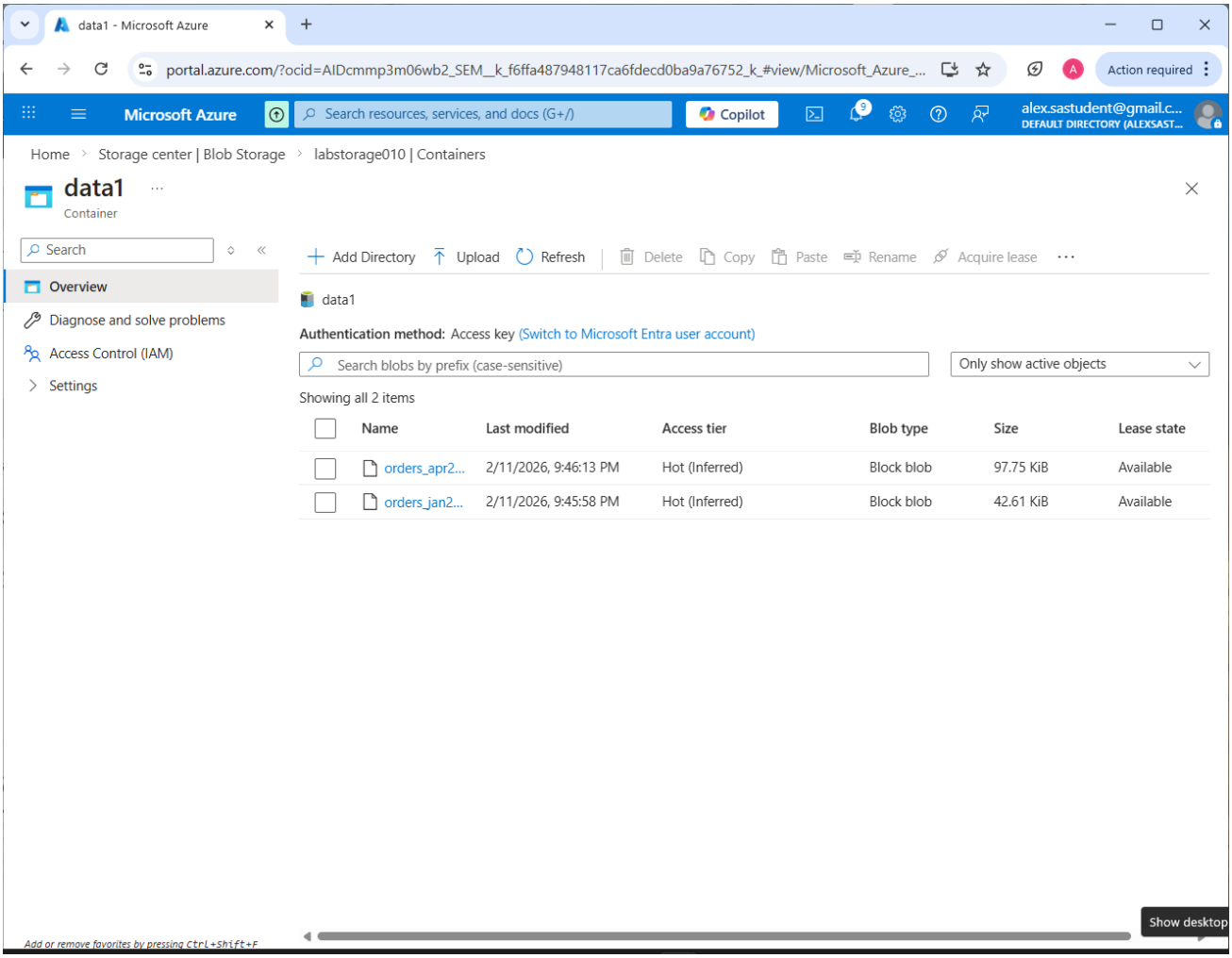

labstorage010_1770799327171
portal.azure.com/?ocid=AIDcmmp3m06wb2_SEM__k_f6ffa487948117ca6fdecd0ba9a76752_k_#view/HubsExten...
Action required
Microsoft Azure
Search resources, services, and docs (G+/)
Copilot
alex.sastudent@gmail.c...
DEFAULT DIRECTORY (ALEXSAST...
Home
labstorage010_1770799327171 | Overview
Deployment
Search
Overview
Inputs
Outputs
Template
Delete
Cancel
Redeploy
Download
Refresh
Your deployment is complete
Deployment name: labstorage01...
Subscription: Azure subscription 1
Resource group: rg-lab
Start time: 2/11/2026, 9:42:21 PM
Correlation ID: 19959987-8b5a-4004-9460-0
Deployment details
Next steps
Go to resource
Give feedback
Tell us about your experience with deployment
Cost Management
Get notified to stay within your budget and prevent unexpected charges on your bill.
Set up cost alerts >
Microsoft Defender for Cloud
Secure your apps and infrastructure
Go to Microsoft Defender for Cloud >
Free Microsoft tutorials
Start learning today >
Work with an expert
Azure experts are service provider partners who can help manage your assets on Azure and be your first line of support.
Find an Azure expert >
Add or remove favorites by pressing Ctrl+Shift+F


New container - Microsoft Azur
portal.azure.com/?ocid=AIDcmmp3m06wb2_SEM__k_f6ffa487948117ca6fdecd0ba9a76752_k_#@alexsastudent...
Action required
Microsoft Azure
Search resources, services, and docs (G+/)
Copilot
alex.sastudent@gmail.c...
DEFAULT DIRECTORY (ALEXSAST...
Home
Storage center | Blob Storage
labstorage010
labstorage010 | Containers
Storage account
Search
Overview
Activity log
Tags
Diagnose and solve problems
Access Control (IAM)
Data migration
Events
Storage browser
Partner solutions
Resource visualizer
Data storage
Containers
File shares
Queues
Tables
Security + networking
Data management
Settings
Monitoring
Monitoring (classic)
Add or remove favorites by pressing Ctrl+Shift+F
Add container
Upload
Refresh
Delete
Search containers by prefix
Showing all 1 items
Name
Last modified
$logs
2/11/2026, 9:
New container
Name *
data
Anonymous access level
Private (no anonymous access)
The access level is set to private because anonymous access is disabled on this storage account.
Advanced
Create
Give feedback



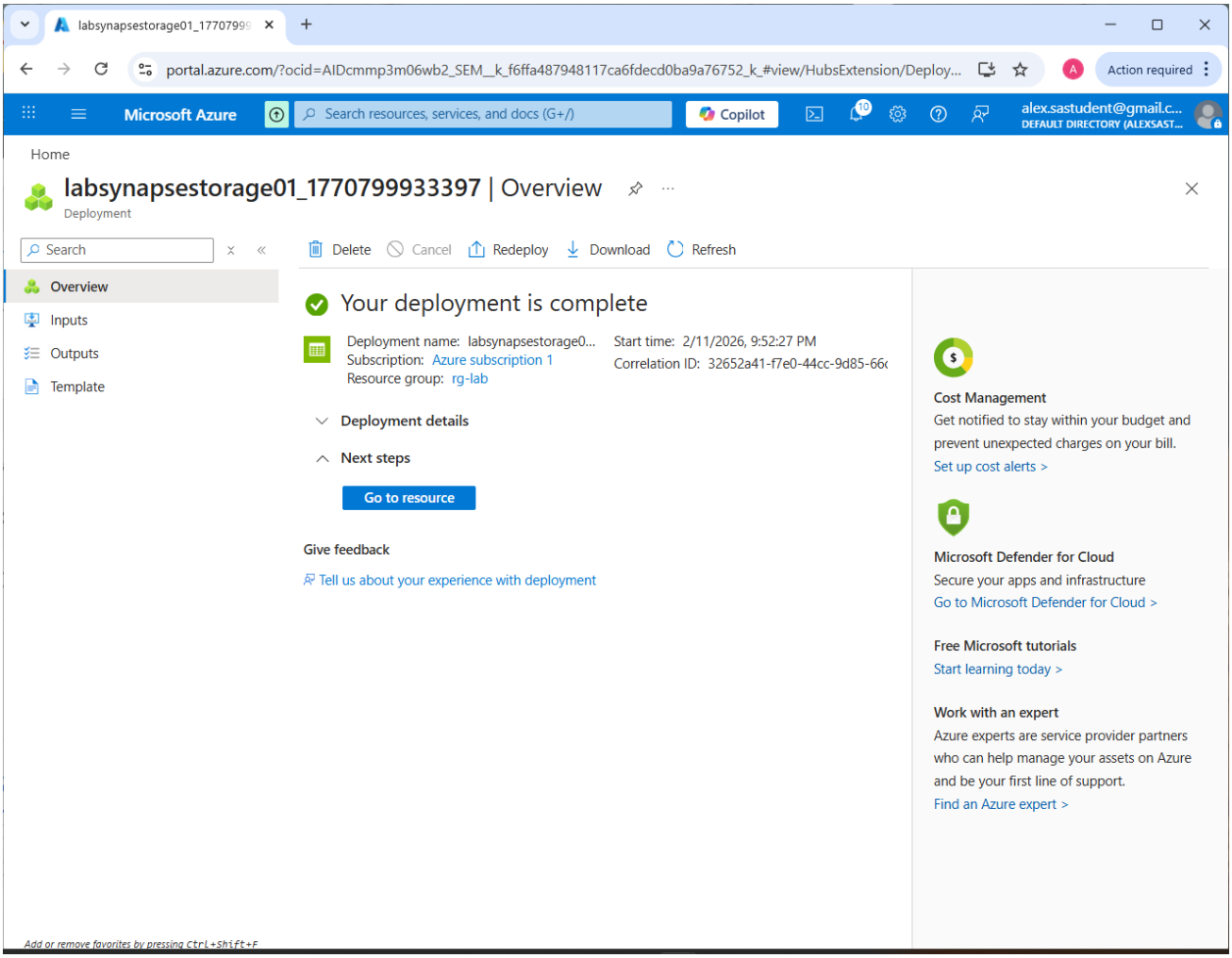
Activities 3.2: Upload CSV files from Azure Portal

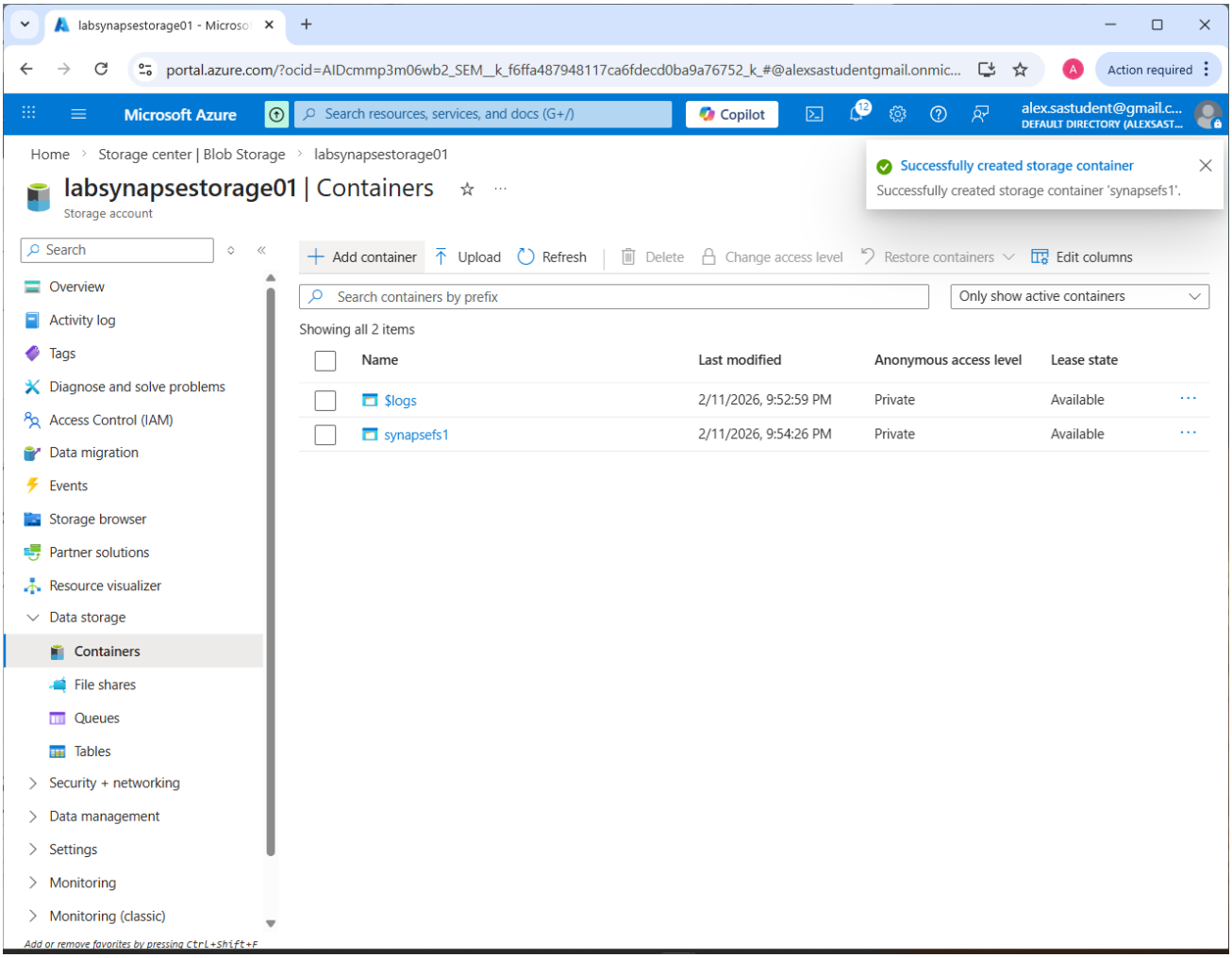
The CSV files were uploaded to an Azure Blob container.



Activities 3.3: Ingest Data into Synapse

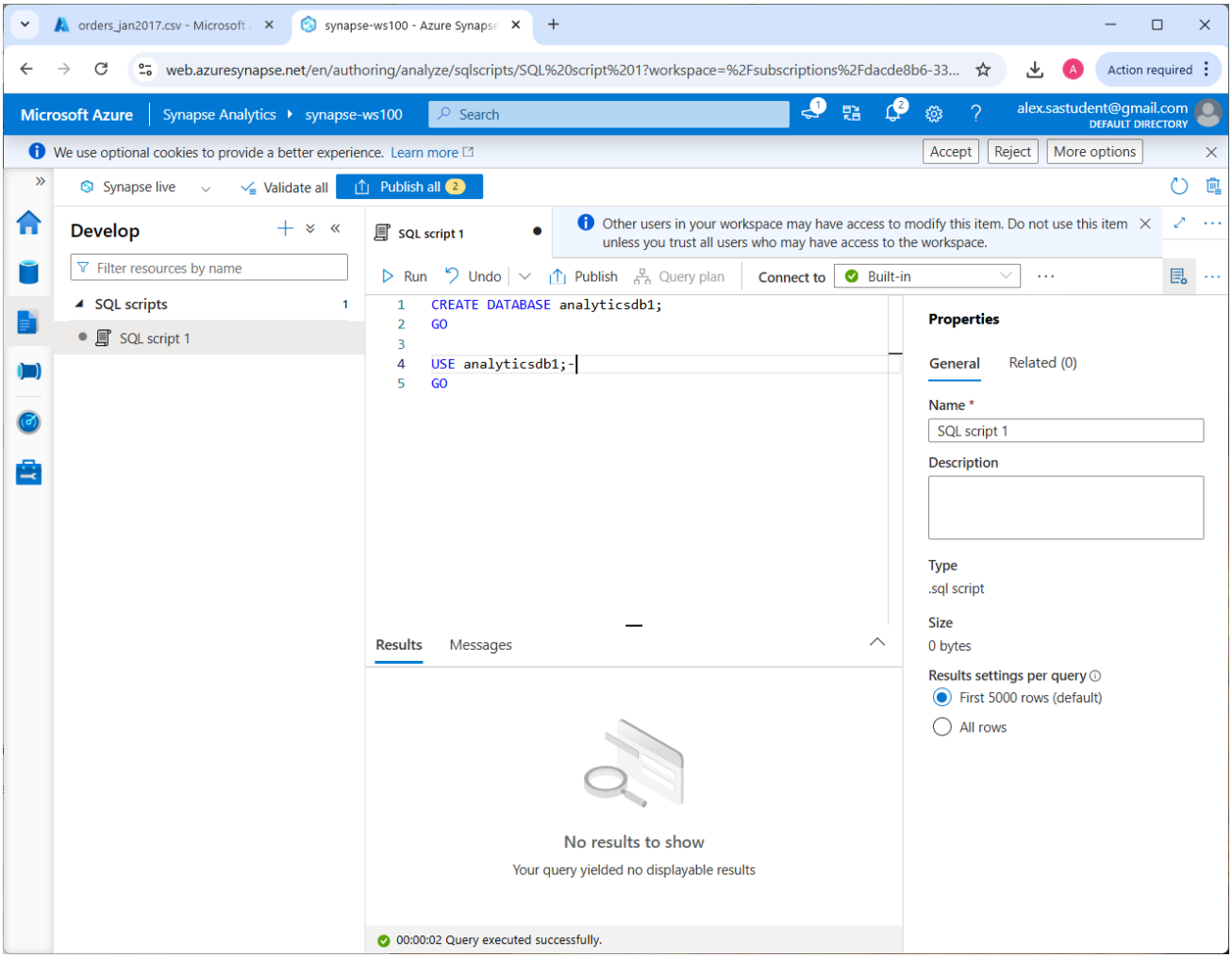
A Linked Service was configured in Synapse to connect to the storage account, and the Copy Data tool was used to ingest the data into a SQL table.

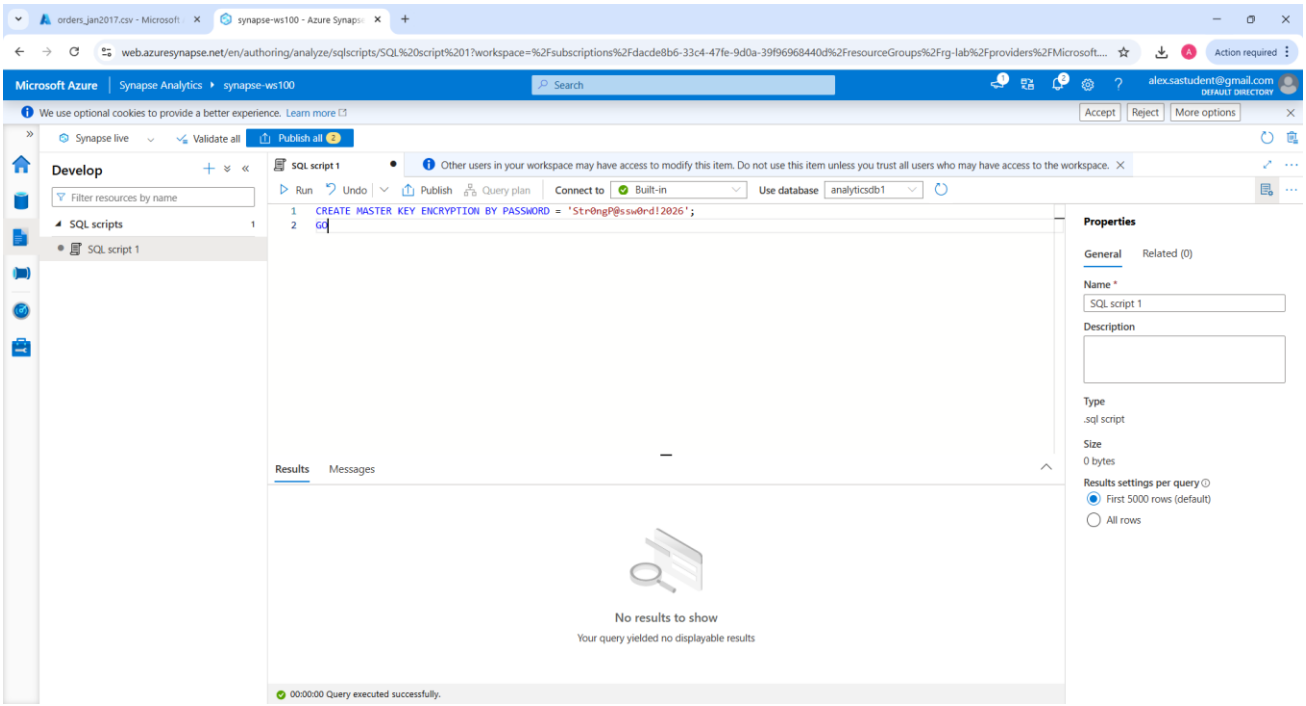

labsynapsestorage01_1770799
portal.azure.com/?ocid=AIDcmmp3m06wb2_SEM__k_f6ffa487948117ca6fdecd0ba9a76752_k_#view/HubsExtension/Deploy...
Action required
Microsoft Azure
Search resources, services, and docs (G+/)
Copilot
alex.sastudent@gmail.c...
DEFAULT DIRECTORY (ALEXSAST...
Home
labsynapsestorage01_1770799933397 | Overview
Deployment
Search
Overview
Inputs
Outputs
Template
Delete
Cancel
Redeploy
Download
Refresh
Your deployment is complete
Deployment name: labsynapsestorage0...
Subscription: Azure subscription 1
Resource group: rg-lab
Start time: 2/11/2026, 9:52:27 PM
Correlation ID: 32652a41-f7e0-44cc-9d85-66
Deployment details
Next steps
Go to resource
Give feedback
Tell us about your experience with deployment
Cost Management
Get notified to stay within your budget and prevent unexpected charges on your bill.
Set up cost alerts >
Microsoft Defender for Cloud
Secure your apps and infrastructure
Go to Microsoft Defender for Cloud >
Free Microsoft tutorials
Start learning today >
Work with an expert
Azure experts are service provider partners who can help manage your assets on Azure and be your first line of support.
Find an Azure expert >
Add or remove favorites by pressing Ctrl+Shift+F

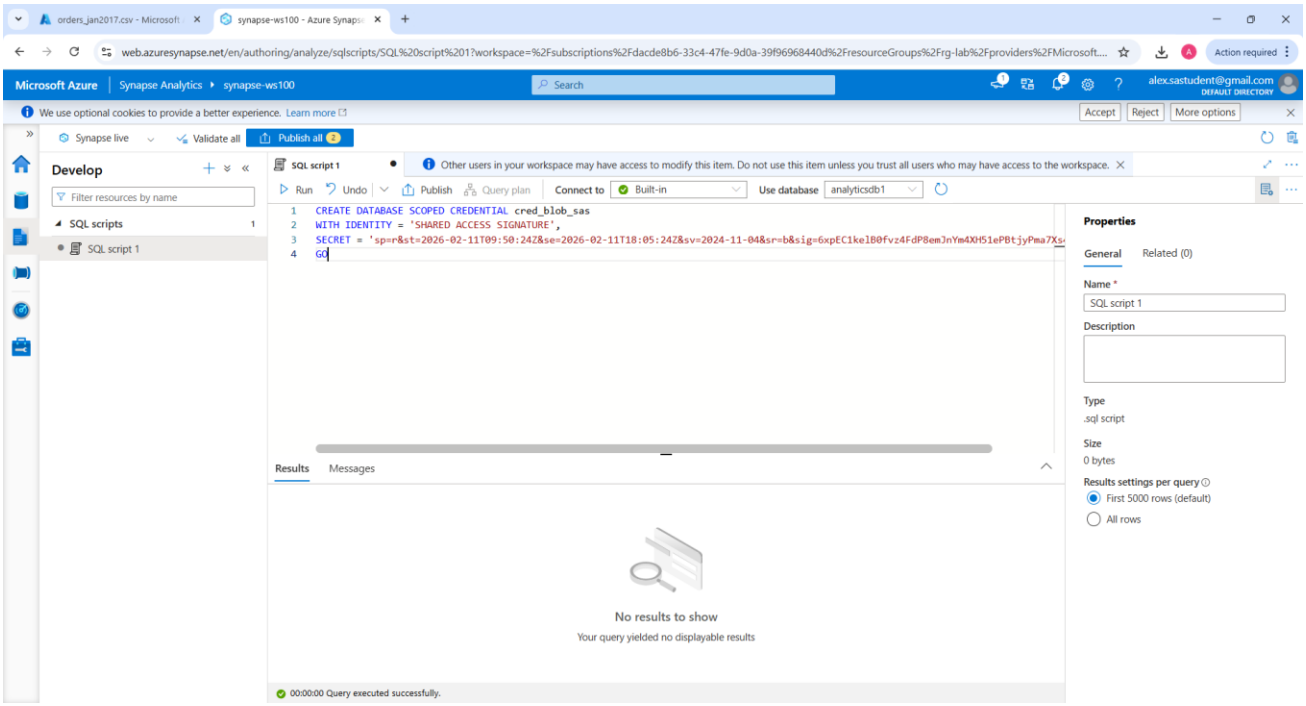


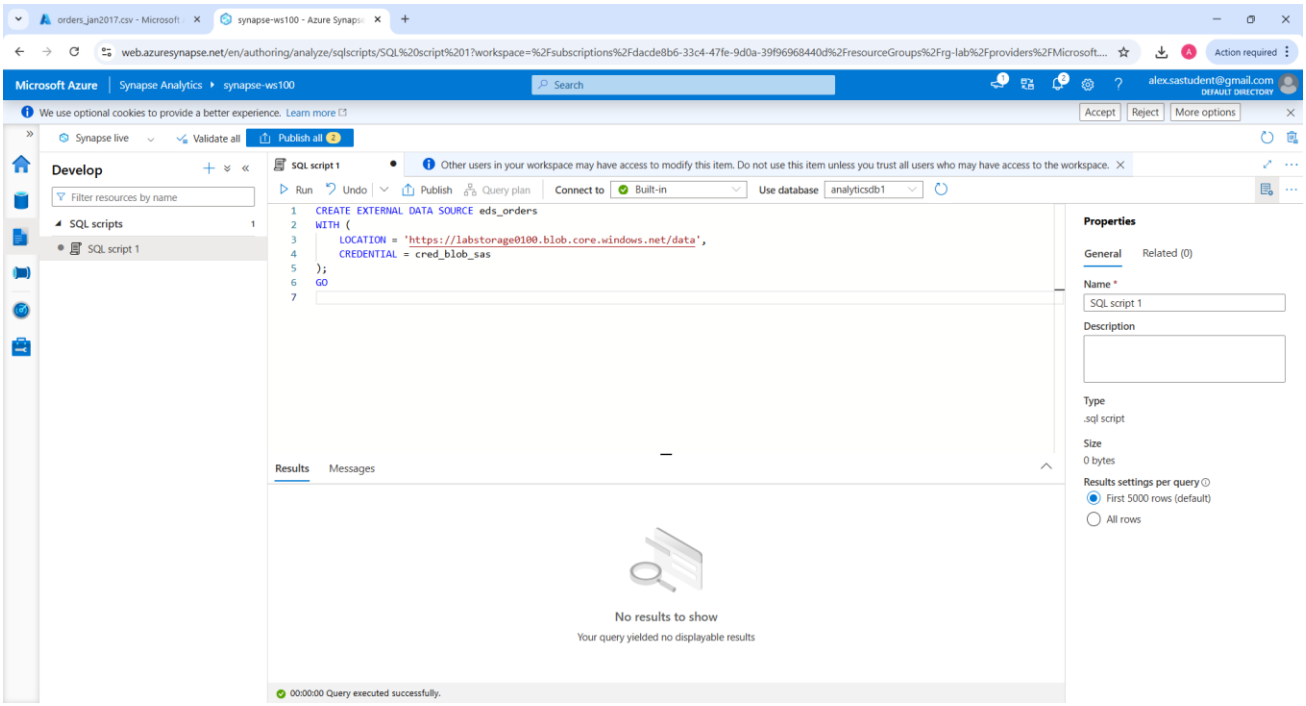
Activities 3.4: Run SQL Queries in Synapse

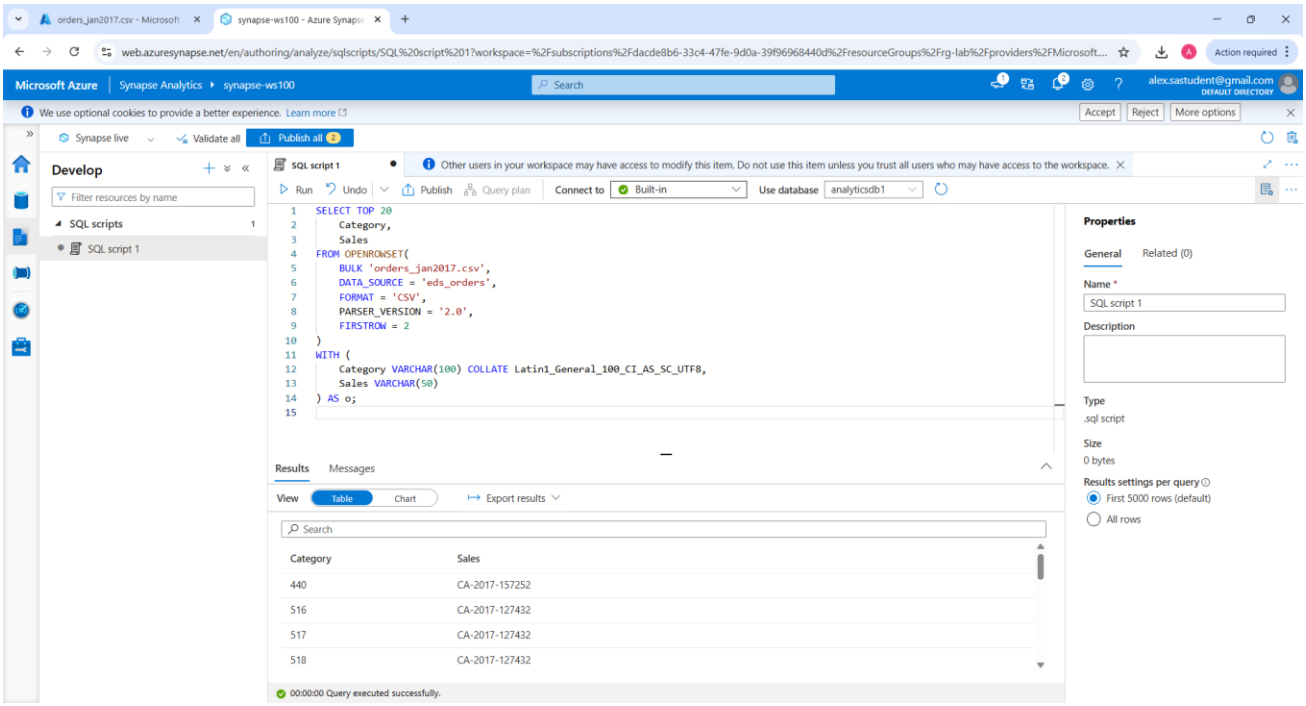
SQL queries similar to those used in Athena were executed in Synapse to validate the data migration. The output confirmed successful ingestion.


orders_jan2017.csv - Microsoft
synapse-ws100 - Azure Synapse
web.azuresynapse.net/en/authoring/analyze/sqlscripts/SQL%20script%201?workspace=%2Fsubscriptions%2Fdacde8b6-33...
Action required
Microsoft Azure
Synapse Analytics
synapse-ws100
Search
alex.sastudent@gmail.com
DEFAULT DIRECTORY
We use optional cookies to provide a better experience. Learn more
Accept
Reject
More options
Synapse live
Validate all
Publish all 2
Develop
Filter resources by name
SQL scripts 1
SQL script 1
Other users in your workspace may have access to modify this item. Do not use this item unless you trust all users who may have access to the workspace.
Run
Undo
Publish
Query plan
Connect to
Built-in
1 CREATE DATABASE analyticsdb1;
2 GO
3
4 USE analyticsdb1;
5 GO
Properties
General
Related (0)
Name *
SQL script 1
Description
Type
.sql script
Size
0 bytes
Results settings per query
First 5000 rows (default)
All rows
Results
Messages
No results to show
Your query yielded no displayable results
00:00:02 Query executed successfully.


Microsoft Azure
Synapse Analytics
synapse-ws100
Develop
SQL script 1
Use database
analyticsdb1
1 CREATE MASTER KEY ENCRYPTION BY PASSWORD = 'Str0ngP@ssw0rd!2026';
2 GO
No results to show
Your query yielded no displayable results
00:00:00 Query executed successfully.


Microsoft Azure | Synapse Analytics | synapse-ws100
Develop
SQL scripts
SQL script 1
CREATE DATABASE SCOPED CREDENTIAL cred_blob_sas
WITH IDENTITY = 'SHARED ACCESS SIGNATURE',
SECRET = 'sp=r&st=2026-02-11T09:50:24Z&se=2026-02-11T18:05:24Z&sv=2024-11-04&sr=b&sig=6xpEC1kelB0fvz4FdP8emJnYm4XH51ePBtjyPma7Xs
GO
Results
Messages
No results to show
Your query yielded no displayable results
00:00:00 Query executed successfully.
Properties
General
Related (0)
Name *
SQL script 1
Description
Type
.sql script
Size
0 bytes
Results settings per query
First 5000 rows (default)
All rows


Microsoft Azure | Synapse Analytics | synapse-ws100
Develop
SQL scripts
SQL script 1
CREATE EXTERNAL DATA SOURCE eds_orders
WITH (
LOCATION = 'https://labstorage0100.blob.core.windows.net/data',
CREDENTIAL = cred_blob_sas
);
GO
Results
Messages
No results to show
Your query yielded no displayable results
00:00:00 Query executed successfully.
Properties
General
Related (0)
Name *
SQL script 1
Description
Type
.sql script
Size
0 bytes
Results settings per query
First 5000 rows (default)
All rows


Microsoft Azure
Synapse Analytics
synapse-ws100
Develop
SQL scripts
SQL script 1
Run
Undo
Publish
Query plan
Connect to
Built-in
Use database
analyticsdb1
SELECT TOP 20
Category,
Sales
FROM OPENROWSET(
BULK 'orders_jan2017.csv',
DATA_SOURCE = 'eds_orders',
FORMAT = 'CSV',
PARSER_VERSION = '2.0',
FIRSTROW = 2
)
WITH (
Category VARCHAR(100) COLLATE Latin1_General_100_CI_AS_SC_UTF8,
Sales VARCHAR(50)
) AS o;
Results
Messages
Table
Chart
Export results
Category
Sales
440
CA-2017-157252
516
CA-2017-127432
517
CA-2017-127432
518
CA-2017-127432
00:00:00 Query executed successfully.
Properties
General
Related (0)
Name *
SQL script 1
Description
Type
.sql script
Size
0 bytes
Results settings per query
First 5000 rows (default)
All rows

Task 4: Final Report on AWS Services and Architecture

Activities 4.1: Report writing

Introduction

In this assessment, an automated cloud-based data pipeline was designed using AWS services to simulate a real-world data engineering workflow. The main objective was to ingest, catalogue, query, visualise, and monitor datasets using cloud-native services instead of traditional manual methods. The solution also included exporting curated results to Microsoft Azure Synapse Analytics to demonstrate multi-cloud interoperability.
This task helped me understand how AWS services can be integrated together to build a scalable and automated data infrastructure using service-based tools. Instead of writing large amounts of custom code, the system was built using managed services such as IAM, S3, Glue, Athena, QuickSight, Lambda, and SNS. This directly aligns with Learning Outcome 3 (LO3), which focuses on designing automated IT infrastructure using user-friendly programming and cloud services.

Overview of AWS Services Used

AWS Identity and Access Management (IAM)

IAM was used to control access to AWS resources securely. An IAM user was created and assigned appropriate permissions to manage S3, Glue, Athena, and other services. This step ensured that all actions performed within the environment were authorised and traceable.
Using IAM helped me understand the importance of role-based access control in cloud environments. Without proper permissions, services such as Glue and Athena cannot interact with S3 correctly.

Amazon S3 (Simple Storage Service)

Amazon S3 acted as the main storage layer of the data pipeline. A bucket was created to store the CSV datasets, and folders were organised using partition-style naming (for example, snapshot_day=2017-01-01).
S3 served as the raw data layer, where files were stored in their original format before being processed. Structuring folders properly was important because Glue used these paths to detect partitions automatically.

AWS Glue

AWS Glue was used for automated data cataloguing. A Glue database (db_yoobee) and a crawler (yoobee_crawler) were configured to scan the S3 bucket and detect schema automatically.
When a new folder was added (for example, snapshot_day=2017-01-04), the crawler was re-run, and new partitions were created without manual schema updates. This demonstrated how automation reduces repetitive work in data engineering.

Amazon Athena

Athena was used as a serverless query engine to run SQL queries directly on data stored in S3. It connected to the Glue Data Catalog to read table metadata.
Queries such as summing sales by category and filtering by snapshot date were executed without creating a database server. Athena follows a pay-per-query model, which makes it cost-efficient for analytics tasks.

Amazon QuickSight

Amazon QuickSight was used for data visualisation. After connecting it to Athena and authorising S3 access, dashboards were created to visualise sales by category and profit by city.
QuickSight uses SPICE to import datasets and generate visual insights without writing additional code. This demonstrates how business intelligence tools can directly integrate with cloud analytics services.

How These Services Work Together (End-to-End Workflow)

The process begins when CSV files are uploaded into the S3 bucket. This upload triggers the Lambda function automatically. If the file extension is valid, SNS sends a notification. Meanwhile, the Glue crawler scans the S3 folder structure and updates the table metadata. Athena then queries the updated table, and QuickSight connects to Athena for visualisation. Finally, selected outputs are transferred to Azure Synapse for additional analysis.

Automation and No-Code Features

One of the most important aspects of this project was automation without heavy coding.
The Glue crawler automatically detected schema and partitions without requiring manual table creation. Athena allowed SQL queries to be executed directly on S3 data without provisioning a database server. QuickSight allowed me to create dashboards using a graphical interface instead of writing additional code.
This approach reduced manual configuration and eliminated the need to provision virtual machines or manage infrastructure. The system mainly relied on managed services provided by AWS.
Through this process, I understood how modern cloud platforms reduce complexity by offering service-based automation rather than traditional server management.

Data Engineering Perspective

From a data engineering perspective, the solution followed a layered approach. Raw data was stored in S3, metadata was managed through Glue, and queries were executed using Athena. Visual insights were generated in QuickSight, while Lambda and SNS handled event-driven monitoring. Seeing these components connected helped me understand how real-world data pipelines are structured.
Partitioned data design improved query efficiency by reducing the data scanned size. This is important in real-world environments where cost optimisation matters.
The pipeline also supports scalability. If new data snapshots are uploaded, the crawler updates partitions automatically, and queries continue to work without redesigning the system.
This is similar to how enterprise data pipelines are designed in real-world environments.

Before this task, I understood these concepts theoretically, but implementing them in a connected workflow made the architecture clearer to me.

Athena vs Azure Synapse SQL – Brief Comparison

During this assessment, both Amazon Athena and Azure Synapse SQL were used for querying data. Athena is fully serverless and requires minimal setup. It directly queries data stored in S3 using the Glue catalog. It is simple and fast for analytics tasks.
Azure Synapse SQL also supports serverless querying, but requires linking storage accounts and creating external tables. It offers stronger integration within the Azure ecosystem.
From my experience, Athena was easier to configure and quicker to use for simple analytical queries. Synapse required additional configuration, such as linking storage accounts and creating external tables. However, Synapse felt more structured for enterprise-level environments. Both platforms operate on a pay-per-query model and support large-scale analytics without managing servers.

Challenges and Solutions

Several challenges were encountered during implementation.
One issue was related to IAM permissions. Initially, Athena could not access the S3 bucket due to missing permissions. This was resolved by attaching the correct policies to the IAM role used by Glue and Athena.
Another issue occurred when the Glue crawler did not detect new partitions after adding a new folder. This happened because the folder structure was not consistent with partition naming. After correcting the naming format (snapshot_day=YYYY-MM-DD), the crawler updated successfully.
QuickSight initially showed S3 access errors. This was resolved by authorising QuickSight to access the S3 bucket in security settings.
Additionally, Athena returned empty results when querying a newly added dataset. After re-running the crawler and refreshing metadata, the issue was fixed.
These challenges improved my understanding of cloud permissions, schema management, and service integration.
At one point, I assumed the table had updated correctly because the crawler completed successfully. However, when I ran the Athena query, the results did not reflect the new data. After reviewing the partition structure, I realised the folder naming format was slightly inconsistent. Correcting the folder naming and re-running the crawler resolved the issue.

Summary and Learning Outcome Reflection

Overall, this project gave me practical experience in building a serverless data pipeline using AWS services. Each service played a specific role, and together they formed a scalable and cost-effective architecture.
Through this implementation, I gained practical experience in configuring IAM, designing S3 storage structures, automating schema detection with Glue, querying with Athena, and visualising data using QuickSight. The addition of Lambda and SNS further demonstrated event-driven automation.
This project helped me to be familiar with designing and automating infrastructure using cloud-based services rather than manual configuration. It reflects real-world cloud engineering practices and strengthened my understanding of multi-cloud data architecture.